目　录

第九章 线 性 群

§9.1 基底的变换

在矢量空间V中，矢量ξ的坐标依赖于V的基底的选取(见§7.8)，因此，基底的任意变化将引起ξ的坐标的变化．例如，在实平面$\mathbf{R}^2$中，矢量$\beta=4\varepsilon_1+2\varepsilon_2$，按照定义它关于由单位矢量$\varepsilon_1$和$\varepsilon_2$组成的基底的坐标是(4,2)．矢量

$$\alpha_1=2\varepsilon_1,\quad \alpha_2=\varepsilon_1+\varepsilon_2 \tag{1}$$

也可以构成一组基底，关于这组基底，β可表示成$\beta=\alpha_1+2\alpha_2$．系数1和2是β关于这组新基底的坐标(也就是关于图1所表示的斜角坐标系的坐标)．

更一般地，任意矢量ξ关于新基底α_1,α_2的坐标x_1^*,x_2^*可以从ξ的"老"坐标x_1，x_2按下面方法求得．按照定义(§7.8)，这两组坐标是矢量ξ关于两组基底表达式的系数

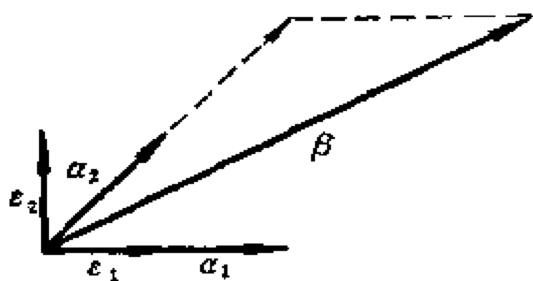


图 1

$$\xi=x_1\varepsilon_1+x_2\varepsilon_2,\quad \xi=x_1^*\alpha_1+x_2^*\alpha_2.$$

解矢量方程(1)，我们求出ε_1和ε_2

$$\varepsilon_1=\frac{1}{2}\alpha_1,\quad \varepsilon_2=\alpha_2-\frac{1}{2}\alpha_1.$$

把ε_1和ε_2的值代入ξ的第一个表达式中，我们得到

$$\xi=x_1\left(\frac{1}{2}\alpha_1\right)+x_2\left(\alpha_2-\frac{1}{2}\alpha_1\right)=\frac{1}{2}(x_1-x_2)\alpha_1+x_2\alpha_2.$$

因此 ξ 的新坐标由线性齐次方程

$$x_1^* = \frac{1}{2}(x_1 - x_2), \quad x_2^* = x_2 \tag{2}$$

给出. 反过来, 老坐标可以通过新坐标表示为

$$x_1 = 2x_1^* + x_2^*, \quad x_2 = x_2^*.$$

在 n 维空间中有类似的关系式. 如果 $\alpha_1, \cdots, \alpha_n$ 是一组给定的基底, 这些矢量是按一定次序排列的, 而 $\alpha_1^*, \cdots, \alpha_n^*$ 是一组新的(有序)基底, 那么新基底的每个矢量 α_i^* 可以表示为老基底矢量的线性组合:

$$\alpha_i^* = p_{i1}\alpha_1 + \cdots + p_{in}\alpha_n = \sum_{j=1}^{n} p_{ij}\alpha_j, \quad i = 1, \cdots, n. \tag{3}$$

表达式(3)可以形式地写成矩阵方程 $\alpha^* = \alpha P^T$, 这里 P^T 是 P 的转置矩阵, $\alpha = (\alpha_1, \cdots, \alpha_n)$, $\alpha^* = (\alpha_1^*, \cdots, \alpha_n^*)$.

这些表达式的系数矩阵 $P = (p_{ij})$ 的第 i 行元素是矢量 α_i^* 的老坐标 $(p_{i1}, \cdots, p_{in})$. 因为矢量 $\alpha_1^*, \cdots, \alpha_n^*$ 构成一组基底, 所以 P 的所有行是线性无关的, 因此 P 是非奇异的(§ 8.6 定理 9). 反过来, 如果 $P = (p_{ij})$ 是任意非奇异矩阵, $\alpha_1, \cdots, \alpha_n$ 是 V 的任意一组基底, 那么用矩阵 P 按公式(3)所确定的矢量 $\alpha_1^*, \cdots, \alpha_n^*$ 是线性无关的, 因此构成 V 的一组新基底. 这就证明了

定理 1 *如果 $\alpha_1, \cdots, \alpha_n$ 是矢量空间 V 的一组基底, 那么对每个非奇异矩阵 $P = (p_{ij})$, n 个矢量 $\alpha_i^* = \sum p_{ij}\alpha_j (i = 1, \cdots, n)$ 构成 V 的一组新基底, 并且 V 的每组基底可以按这种方法恰由一个非奇异矩阵 P 得到.*

我们还可以用新基底表示老基底, 其方程式是 $\alpha_k = \sum q_{ki}\alpha_i^*$, 其中系数矩阵为 $Q = (q_{ki})$. 如果把用 $\alpha_1, \cdots, \alpha_n$ 表示的 α_i^* 的值代入这个方程式中, 我们就得到

$$\alpha_k=\sum_i q_{ki}\Big(\sum_j p_{ij}\alpha_j\Big)=\sum_j\Big(\sum_i q_{ki}p_{ij}\Big)\alpha_j.$$

可是,α_1, …, α_k 这组矢量用它们本身来表示的话，只有一种表达式，即 $\alpha_k=\alpha_k$. 因此这里 α_j 的系数 $\sum q_{ki}p_{ij}$ 一定是 0 或 1, 它取决于 $k\neq j$ 或 $k=j$. 由于这些系数正好是乘积矩阵 QP 的(k, j)位置上的元素,因此 $QP=I$, 所以 $Q=P^{-1}$ 是 P 的逆矩阵.

与此平行的关于坐标变换的结果如下所述:

定理 2 如果矢量空间 V 的基底 $\alpha_1, \cdots, \alpha_n$ 变换成一组新基底 $\alpha_1^*, \cdots, \alpha_n^*$, α_i^* 表示成形式 $\alpha_i^*=\sum_j p_{ij}\alpha_j$, 那么任意矢量 ξ 关于老基底 $\alpha_1, \cdots, \alpha_n$ 的坐标 $x_1, \cdots, x_n$ 通过齐次线性方程组

$$x_j=x_1^*p_{1j}+\cdots+x_n^*p_{nj}=\sum_{i=1}^n x_i^*p_{ij} \tag{4}$$

可以确定 ξ 关于基底 $\alpha_1^*, \cdots, \alpha_n^*$ 的新坐标 $x_1^*, \cdots, x_n^*$.

证明 按照定义(§ 7.8), ξ 关于基底 $\alpha_1^*, \cdots, \alpha_n^*$ 的坐标 $x_1^*, \cdots, x_n^*$, 是把 ξ 看作 $\alpha_1^*, \cdots, \alpha_n^*$ 的线性组合 $\xi=\sum x_i^*\alpha_i^*$ 时，表达式中的系数. 把关于 α_i^* 的公式(3)代入这个表达式中,便得出

$$\xi=\sum_i x_i^*\Big(\sum_j p_{ij}\alpha_j\Big)=\sum_j\Big(\sum_i x_i^*p_{ij}\Big)\alpha_j.$$

这里每个 α_j 的系数是 ξ 的老坐标 x_j, 因此方程组(4)成立.

方程组(4)还可以写成矩阵形式 $X=X^*P$, 这里 $X=(x_1, \cdots, x_n)$是老坐标的行矩阵, $X^*=(x_1^*, \cdots, x_n^*)$是新坐标的行矩阵. 因为 $\alpha_1, \cdots, \alpha_n$ 和 $\alpha_1^*, \cdots, \alpha_n^*$ 都是基底，所以 P 是非奇异的，并且可以利用 X 表示 X^* 为 $X^*=XP^{-1}$.

如果我们把这个矩阵方程同前面已提到的(3)式的矩阵公式 $\alpha^*=\alpha P^T$ 进行比较,便可得到有趣的关系

$$\text{基底: } \alpha^*=\alpha P^T, \quad \text{坐标: } X^*=XP^{-1}. \tag{5}$$

第二个方程中的矩阵 P^{-1} 是第一个方程中的矩阵 P' 的转置逆矩阵.（有时把这些叙述概括成：坐标变换是基底变换的转置逆变换.）

习 题

1. 设 T 是把通常的单位矢量 ε_i（在 V_2 或 V_3 中）变换到下面指定的矢量 α_i. 求出相应的用老坐标表示新坐标的方程，和用新坐标表示老坐标的方程. 对情况(a)和(b)画出图来.

(a) $\alpha_1=(1,1)$, $\alpha_2=(1,-1)$.

(b) $\alpha_1=(2,3)$, $\alpha_2=(-2,-1)$.

(c) $\alpha_1=(1,1,0)$, $\alpha_2=(1,0,1)$, $\alpha_3=(0,1,1)$.

(d) $\alpha_1=(i,1,i)$, $\alpha_2=(0,1,i)$, $\alpha_3=(0,i,1)$，这里 $i^2=-1$.

2. 如果新基底 $\alpha_1^*,\cdots,\alpha_n^*$ 是通过形为 $\alpha_i=\sum_j q_{ij}\alpha_j^*\ (i=1,\cdots,n)$ 的方程组间接给出，算出相应的坐标变换的方程.

3. 给出平面上坐标轴旋转 θ 角的坐标变换的方程.

§9.2 相似矩阵与特征矢量

矢量空间 V 的线性变换 $T:V\to V$，可以用各种不同的矩阵来表示，这些矩阵依赖于 V 的基底(坐标系)的选择. 比如，在平面上，由 $\varepsilon_1\longmapsto 3\varepsilon_1$, $\varepsilon_2\longmapsto -\varepsilon_1+2\varepsilon_2$ 定义的变换，它在 $\mathbf{R}^2$ 的普通坐标系中可用矩阵 A 表示出来，矩阵 A 的行是 ε_1 和 ε_2 的变换式的坐标，如下所示：

$$A=\begin{pmatrix}3&0\\-1&2\end{pmatrix}.$$

但是对于 §9.1 中所讨论的新基底 $\alpha_1=2\varepsilon_1$, $\alpha_2=\varepsilon_1+\varepsilon_2$, 上述变换便是 $\alpha_1\longmapsto 3\alpha_1$, $\alpha_2\longmapsto 2\alpha_2$；因此它就可以用比较简单的对角矩阵

$$D=\begin{pmatrix}3&0\\0&2\end{pmatrix}$$

来表示. 我们称这样两个矩阵 A 和 D 是相似的.

为了推广这个结果，让我们回想一下，矩阵是怎样表示变换的. 取矢量空间 V 的任意一组(有序)基底 $\alpha_1, \cdots, \alpha_n$ 和任意线性变换 $T:V\to V$. 那么基底矢量 α_i 在 T 之下的象可以用 § 8.1 公式(9)写成

$$\alpha_i T=\sum_j a_{ij}\alpha_j, \quad A=(a_{ij}). \tag{6}$$

因此，对于基底 $\alpha=\{\alpha_1, \cdots, \alpha_n\}$, T 用 $n\times n$ 矩阵 A 表示. 这个关系也可以通过坐标来表示. 设 $\xi=\sum x_i\alpha_i$ 是 V 的一个矢量，它对于基底 α 的坐标是 n-数组 $X=(x_1, \cdots, x_n)$. 那么象 $\eta=\xi T$ 是

$$\begin{aligned}\xi T&=(\sum x_i\alpha_i)T=\sum x_i(\alpha_i T)\\&=\sum_i\sum_j x_i a_{ij}\alpha_j=\sum_j\left(\sum_i x_i a_{ij}\right)\alpha_j.\end{aligned}$$

而 η 的坐标 y_j 恰好是 α_j 的系数，所以

$$y_j=\sum_i x_i a_{ij},$$

η 的坐标矢量 Y 恰好是矩阵乘积 $Y=XA$. 简明地写出就是

$$\begin{aligned}Y=XA,\ &\text{其中 } X \text{ 是 } \xi \text{ 关于 } \alpha \text{ 的坐标},\\ &Y \text{ 是 } \eta=\xi T \text{ 关于 } \alpha \text{ 的坐标}.\end{aligned} \tag{7}$$

两个等价的命题(6)和(7)都意味着，对于基底 α, 变换 T 可用矩阵 A 来表示.

现在设 $\alpha_1^*, \cdots, \alpha_n^*$ 是另一组基底. 那么根据定理 1, 新基底可以用一个 $n\times n$ 非奇异矩阵 P 通过老基底来表示，如(3)所示；再根据定理 2, ξ 和 ξT 的新坐标可以通过老坐标给出，表示成 $X^*=XP^{-1}$ 和 $Y^*=YP^{-1}$. 那么由(7)有

$$Y^*=YP^{-1}=XAP^{-1}=X^*(PAP^{-1}).$$

因此再根据(7)，在新坐标系下表示变换 T 的矩阵 B 具有形式

PAP^{-1}. 等价关系 $B=PAP^{-1}$ 在形式上很象群中的共轭元素的关系(§6.12). 在矩阵代数中,这是很重要的,称它为相似关系.

定义 元素在域 F 中的两个 $n\times n$ 矩阵 A 和 B(在 F 上)相似当且仅当在 F 上有一个 $n\times n$ 非奇异矩阵 P 使得 $B=PAP^{-1}$.

上述讨论就证明了

定理3 域 F 上的两个 $n\times n$ 矩阵 A 和 B, 对于(通常)不同的坐标系表示 F 上 n 维矢量空间 V 的同一个线性变换 $T:V\to V$, 当且仅当矩阵 A 和 B 是相似的.

我们还可以更清楚地把这个定理重述如下:

定理3′ 假设对于 V 的基底 $\alpha_1, \cdots, \alpha_n$, 线性变换 $T:V\to V$ 用矩阵 A 来表示,设 $P=(p_{ij})$ 是非奇异矩阵, $\alpha_i{}^*=\sum_j p_{ij}\alpha_j$ 是 V 的相应的新基底,那么对于新基底, T 就用矩阵 PAP^{-1} 来表示.

对角矩阵的代数运算特别容易:任意两个对角矩阵相加(或相乘),只是把相应的对角线元素相加(或相乘). 由于这个以及其他理由,考查什么样的矩阵同对角矩阵相似,考查哪些成对的对角矩阵彼此是相似的, 这是非常重要的. 这些问题的回答包含了特征矢量和特征根的概念,这两个概念也称为本征矢量和本征值.

定义 线性变换 $T:V\to V$ 的特征矢量是 V 中满足条件 $\xi T=c\xi$ 的一个非零矢量 ξ, 这里 c 是某一标量; T 的特征值是满足 $\xi T=c\xi$ 的标量 c, 这里 ξ 是某一非零矢量. 相应地,方阵 A 的特征矢量和特征值是满足 $XA=cX$ 的矢量 $X=(x_1, \cdots, x_n)$ 和标量 c. T(或 T_A)的所有特征值的集合称为 T 的谱.

这样, T 的每个特征矢量 ξ 确定一个特征值 c, 并且每个特征值至少属于一个特征矢量. 因为相似矩阵对应着不同基底下的同一个线性变换,所以相似矩阵具有同样的特征值. 显然,如果矢量 $X\neq 0$ 对某一标量 c 满足 $XA=cX$, n 维矢量 X 就是 $n\times n$ 矩阵 A

的特征矢量．如果矩阵 $B=PAP^{-1}$ 相似于 A，那么 $(XP^{-1})B=XP^{-1}PAP^{-1}=c(XP^{-1})$，所以 XP^{-1} 是 B 的属于同一特征值 c 的特征矢量．还应注意，特征矢量乘上任意非零标量还是特征矢量．

特征矢量与对角矩阵之间的联系由下面定理给出．

定理 4 一个 $n\times n$ 矩阵 A 与一个对角矩阵 D 相似当且仅当 A 的特征矢量张成 F^n；如果 A 与 D 相似，那么 A 的特征值就是 D 的对角线元素．

特别是，这个定理意味着，对角矩阵的特征值是对角线上的元素．

证明 首先假定矩阵 A 与对角矩阵 D 相似，D 的对角线元素是 $d_1,\cdots,d_n$．那么单位矢量 $\varepsilon_1=(1,0,\cdots,0),\cdots,\varepsilon_n=(0,\cdots,0,1)$ 是 D 的特征矢量，这是因为 $\varepsilon_1 D=d_1\varepsilon_1,\cdots,\varepsilon_n D=d_n\varepsilon_n$．还有，对角线元素 $d_1,\cdots,d_n$ 是 D 的相应的特征值，因此也是 A 的特征值．$d_1,\cdots,d_n$ 是唯一的一组特征值，因为设 $X=(x_1,\cdots,x_n)\neq 0$ 是 D 的任意特征矢量，那么对某一适当的特征值 c 有 $XD=cX$．而 $XD=(d_1x_1,\cdots,d_nx_n)$，所以对所有的 i 有 $d_ix_i=cx_i$．因为有某个 $x_i\neq 0$，所以就证明了对这个 i，有 $d_i=c$，于是特征值 c 确实是某个 d_i．

反过来，假设 A 存在足够的特征矢量张成整个空间 F^n，T_A 是 F^n 上相应的线性变换，那么（§7.4 定理 4 的推论 2），我们可以取出特征矢量的一个子集合 $\beta_1,\cdots,\beta_n$，它构成 F^n 的一组基底．因为每个 β_i 是特征矢量，所以有 $\beta_1T_A=c_1\beta_1,\cdots,\beta_nT_A=c_n\beta_n$，其中 $c_1,\cdots,c_n$ 是一组特征值．因此，对于基底 $\beta_1,\cdots,\beta_n$，T_A 可以用对角矩阵 D 象公式(6)那样表示，这里 D 的对角线元素是 $c_1,\cdots,c_n$，所以 A 与这个矩阵 D 相似．

推论 如果矩阵 P 的行是 $n\times n$ 矩阵 A 的 n 个线性无关的特征矢量，那么 P 是非奇异的，并且 PAP^{-1} 是对角矩阵．

证明 我们给出 n 个线性无关的 n 维矢量 $X_1,\cdots,X_n$，它们是

A 的特征矢量，所以对特征值 $c_1, \cdots, c_n$，有 $X_iA=c_iX_i$，$i=1, \cdots, n$. 以 $X_1, \cdots, X_n$ 为行的矩阵 P 是非奇异的，因为它所有的行是线性无关的. 根据分块矩阵乘法法则有

$$\begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} A = \begin{pmatrix} c_1X_1 \\ \vdots \\ c_nX_n \end{pmatrix} = \begin{pmatrix} c_1 & & 0 \\ & \ddots & \\ 0 & & c_n \end{pmatrix} \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}. \tag{8}$$

这就表明 $PA=DP$，因此 $PAP^{-1}=D$，这里 D 是以 $c_1, \cdots, c_n$ 为对角线元素的对角矩阵. 事实上，矩阵 P 恰好是定理 4 的直接证明中对基底变换所需要的矩阵. 证毕

另一方面，存在着与任意对角矩阵都不相似的矩阵(参看下面习题 5).

为了明显地构造出与已知矩阵相似的对角矩阵(如果它存在的话)，我们要找出特征值和特征矢量. 根据下面的考虑，特征值和特征矢量的求法可以大大简化.

如果标量 λ 是 $n\times n$ 矩阵 A 的特征值，I 是 $n\times n$ 单位矩阵，那么 $XA=\lambda X=\lambda XI$，因此对某个非零 n 维矢量 X 有 $X(A-\lambda I)=O$. 于是以 $A-\lambda I$ 为系数矩阵的 n 个齐次线性方程组有非平凡解；因此根据 § 8.6 定理 9 的推论 1，我们有

定理 5 标量 λ 是矩阵 A 的特征值当且仅当矩阵 $A-\lambda I$ 是奇异的.

例如，不难看出，2×2 矩阵

$$A-\lambda I=\begin{pmatrix} a_{11}-\lambda & a_{12} \\ a_{21} & a_{22}-\lambda \end{pmatrix} \tag{9}$$

是奇异的当且仅当

$$\lambda^2-(a_{11}+a_{22})\lambda+a_{11}a_{22}-a_{12}a_{21}=0. \tag{10}$$

(这只表明 $A-\lambda I$ 的行列式等于零.)因此，我们通过求解这个方程求出所有的特征值. 而且，对每个根 λ 至少有一个特征矢量，这

可以通过求解方程组

$$x_1a_{11}+x_2a_{21}=\lambda x_1,$$
$$x_1a_{12}+x_2a_{22}=\lambda x_2$$

而得到.

例 求与矩阵$\begin{pmatrix}-3 & 4\\ 2 & -1\end{pmatrix}$相似的对角矩阵.

多项式(10)是$\lambda^2+4\lambda-5$. 这个多项式的根是1和-5; 因此特征矢量满足齐次方程组

$$\begin{aligned}-3x+2y&=x\\ 4x-y&=y,\end{aligned}\quad 或 \quad \begin{aligned}-3x+2y&=-5x\\ 4x-y&=-5y.\end{aligned}$$

解这两个方程组, 我们得到特征矢量$(1, 2)$和$(1, -1)$. 用这两个矢量作新的基底, 上述变换的矩阵就呈现对角形. 根据定理3′, 新的对角矩阵可以写成矩阵的乘积

$$\begin{pmatrix}1 & 2\\ 1 & -1\end{pmatrix}\begin{pmatrix}-3 & 4\\ 2 & -1\end{pmatrix}\begin{pmatrix}1 & 2\\ 1 & -1\end{pmatrix}^{-1}=\begin{pmatrix}1 & 0\\ 0 & -5\end{pmatrix}.$$

习　　题

1. 证明: 方程$2x'=(1+b)x+(1-b)y$, $2y'=(1-b)x+(1+b)y$, 表示一个关于通过原点与x轴成45°角的直线的压缩. 计算这个变换的特征值和特征矢量, 并说明它们的几何意义.

2. 计算下列在复数域上的矩阵的特征值和特征矢量:

(a) $\begin{pmatrix}2 & 4\\ 5 & 3\end{pmatrix}$,　　(b) $\begin{pmatrix}3 & 2\\ -2 & 3\end{pmatrix}$,

(c) $\begin{pmatrix}1 & 2\\ 2 & -2\end{pmatrix}$,　　(d) $\begin{pmatrix}-1 & 2i\\ -2i & 2\end{pmatrix}$.

3. 对习题2所给出的每个矩阵A, 如果可能的话, 求出非奇异矩阵P, 使得PAP^{-1}是对角矩阵.

4. (a) 求出表示转过θ角的平面旋转的矩阵的复特征值.

(b) 证明: 表示转过θ角$(0<\theta<\pi)$的平面旋转的矩阵不能同任意实

对角矩阵相似.

5. 证明：当 $c\neq 0$, 矩阵 $\begin{pmatrix}1 & c\\ 0 & 1\end{pmatrix}$ 不能同任意实的或复的对角矩阵相似. 从几何上说明这个结果.

6. 证明：2×2 矩阵 A 的特征矢量的斜率 γ 满足二次方程 $a_{21}\gamma^2+(a_{11}-a_{22})\gamma-a_{12}=0$.

7. 证明：属于已知矩阵的固定特征值的所有特征矢量的集合构成一个子空间，这时假定 $\mathbf{0}$ 包含在这些特征矢量中.

8. 证明：非标量矩阵的任意 2×2 实对称矩阵有两个不同的实特征矢量.

9. (a) 证明：两个 $m\times n$ 矩阵 A 和 B 是等价的当且仅当它们对于 m 维矢量空间 V 与 n 维矢量空间 W 的两组不同基底表示同一个从 V 到 W 的线性变换 $T:V\to W$.

(b) 按照这种看法，解释 § 8.9 定理 18.

*10. 设 A 和 B 都与对角矩阵相似. 证明：$AB=BA$ 当且仅当 A 与 B 具有共同的特征矢量基底(Frobenius).

*11. (a) 证明：如果 $A=\begin{pmatrix}a & b\\ c & d\end{pmatrix}$ 与一个正交矩阵相似，那么 $ad-bc=\pm 1$. (正交矩阵的定义见 § 9.4.)

(b) 证明：如果 $ad-bc=1$，那么 A 与正交矩阵相似当且仅当 $A=\pm I$，或 $-2<a+d<2$.

(c) 证明：如果 $ad-bc=-1$，那么 A 与正交矩阵相似当且仅当 $a+d=0$.

§ 9.3 全线性群与仿射群

n 维矢量空间 F^n 的所有非奇异线性变换构成一个群，因为这样的变换的乘积和逆变换还是线性的和非奇异的(§ 8.6 定理 9). 这个群称为全线性群 $L_n=L_n(F)$. 在线性变换同矩阵的一一对应中，线性变换的乘积对应着矩阵的乘积，所以全线性群与元素属于域 F 的所有 $n\times n$ 非奇异矩阵构成的群同构.

全体平移构成另一重要的群．平面上的平移是把平面上的所有点沿着某一指定的方向移动同样一段距离．移动的距离和方向可以用矢量 κ 表示，κ 具有适当的大小和方向，那么平移把每个矢量 ξ 的端点移到矢量 $\xi+\kappa$ 的端点．在空间 F^n 中，平移是变换 $\xi\longmapsto\xi+\kappa$，其中 κ 是固定矢量．对任意坐标系，平移后的矢量的坐标是 $y_1=x_1+k_1$，…，$y_n=x_n+k_n$，这里 k_i 是矢量 κ 的坐标．平移 $\xi\longmapsto\eta=\xi+\kappa$ 与 $\eta\longmapsto\zeta=\eta+\lambda$ 的乘积是通过代入而得，它是平移 $\xi\longmapsto\zeta=\xi+(\kappa+\lambda)$．这恰好对应于矢量 κ 与 λ 的和．类似地，平移 $\xi\longmapsto\xi+\kappa$ 的逆是 $\eta\longmapsto\eta-\kappa$．于是我们就证明了凯莱定理(§ 6.5 定理 8)的特殊情形：

定理 6 *F^n 的所有平移 $\xi\longmapsto\xi+\kappa$ 构成一个阿贝耳群，这个群与 F^n 的全体矢量 κ 的加法群同构．*

线性变换 T 后面再跟随一个平移，就得到变换

$$\xi\longmapsto\eta=\xi T+\kappa\quad(T\text{ 是线性变换},\kappa\text{ 是固定矢量})\qquad(11)$$

任意一个这种形式的变换称为 F^n 的一个*仿射变换* H．仿射变换包括线性变换(当 $\kappa=0$)和平移(当 $T=I$)．如果仿射变换(11)后面跟随第二个仿射变换 $\eta\longmapsto\eta U+\lambda$，则它们的乘积是

$$\xi\longmapsto(\xi T+\kappa)U+\lambda=\xi(TU)+(\kappa U+\lambda).\qquad(12)$$

其结果还是仿射变换，因为 $\kappa U+\lambda$ 是 F^n 的一个固定矢量．每个平移是一一的，也是映上的，因此它有逆．所以仿射变换(11)是一一映上的当且仅当它的线性部分是一一的．因此仿射变换(11)的逆是仿射变换 $\eta\longmapsto\xi=\eta T^{-1}-\kappa T^{-1}$，这可通过公式(11)解出 ξ 而得到．这就证明了

定理 7 *F^n 的所有非奇异仿射变换的集合构成一个群，称为仿射群 $A_n(F)$．全线性群和平移群是它的子群．*

仿射变换对于基底的方程是什么？线性部分 T 产生矩阵 $A=(a_{ij})$；平移矢量按坐标写成行矢量 $K=(k_1,\cdots,k_n)$．于是仿射变

换把坐标为$X=(x_1,\cdots,x_n)$的矢量变换成坐标为

$$Y=XA+K,\quad y_j=\sum_{i=1}^{n}x_i a_{ij}+k_j\quad(j=1,\cdots,n)\tag{13}$$

的矢量．一个变换是仿射变换当且仅当它对于某一组基底可表示成象(13)式那样的非齐次线性方程组．

变换(13)与$Z=YB+L$的乘积是

$$Z=X(AB)+KB+L\quad(K,L\text{是行矩阵}).\tag{14}$$

这个公式与(12)式相平行．我们从变换(13)按下面方式构造一个$n+1$阶矩阵，即在矩阵A的右边加上一个零列，A的下面加上行矢量K，右下方加上单个元素1：

$$\{Y=XA+K\}\longleftrightarrow\begin{pmatrix}A & O\\ K & 1\end{pmatrix}(O\text{是}n\times1\text{矩阵},K\text{是}1\times n\text{矩阵})\tag{15}$$

对于这样的矩阵，满足同样的乘法法则．由分块矩阵乘法法则(§8.5，(43)式)得到

$$\begin{pmatrix}A & O\\ K & 1\end{pmatrix}\begin{pmatrix}B & O\\ L & 1\end{pmatrix}=\begin{pmatrix}AB+OL & AO+O\cdot1\\ KB+1\cdot L & KO+1\cdot1\end{pmatrix}$$
$$=\begin{pmatrix}AB & O\\ KB+L & 1\end{pmatrix}.\tag{16}$$

等式右边这个结果正是对应于乘积变换(14)的加边矩阵．这就证明了

定理8 n维空间的所有非奇异仿射变换构成的群与所有最后一列为$(0,\cdots,0,1)^T$的$(n+1)\times(n+1)$非奇异矩阵构成的群同构．这个同构通过对应(15)明显地给出．

每个仿射变换$\xi H=\xi T+\kappa$确定唯一的线性变换T．按照(12)式，两个仿射变换的乘积确定相应的线性部分的乘积．这个

对应 $H \longmapsto T$ 把非奇异仿射变换群映上到全线性群，在群论意义下(§ 6.11)，它是一个同态．在任意同态中，映射到单位元素的元素构成的集合是一正规子群；这时，满足 $T=I$ 的仿射变换 H 恰好是一个平移．这就证明了

定理 9 *平移群是仿射群的正规子群．*

方程(13)可以象上述那样解释为点(矢量)的变换，它把每个点 $X=(x_1, \cdots, x_n)$ 变换到同一坐标系中的新点 $Y=(y_1, \cdots, y_n)$．同样我们可以把方程(13)解释为坐标的变换．我们称第一种解释为坐标固定图象移动的变换(即点移动到另一个地方)，称第二种解释为图象固定坐标移动的变换(即点取另一个名字)．

例如，在平面上，方程组

$$y_1=x_1+2, \quad y_2=x_2-1,$$

当把它作为第一种变换解释时，它把整个平面向东平移两个单位，再向南平移一个单位；当把它作为第二种变换解释时，原来坐标网格用一个平行的网格来代替，新的坐标原点是把原来的坐标原点向西移动两个单位再向北移动一个单位而得到．

对所有变换群都可作类似的双重解释．

习　　题

1．(a) 用矩阵表示下列各仿射变换：

H_1: $x'=3x+6y+2, \quad y'=3y-4$;

H_2: $x'=x+y+3, \qquad y'=x-y+5$

(b) 计算乘积 H_1H_2，H_2H_1．

(c) 求 H_1 和 H_2 的逆．

2．证明：满足条件 $ad-bc=1$ 的所有仿射变换 $x'=ax+by+e, y'=cx+dy+f$ 的集合是仿射群 $A_2(F)$ 的一个正规子群．

*3．已知单位圆 $x^2+y^2=1$，证明：平面上每个非奇异仿射变换把这个单位圆变成椭圆或圆．

4. 在下列域F上$n\times n$矩阵的集合中，哪些是全线性群的子群？

(a) 全体标量矩阵 cI.

(b) 全体对角矩阵.

(c) 全体非奇异对角矩阵.

(d) 全体置换矩阵.

(e) 全体单项矩阵.

(f) 全体三角形矩阵.

(g) 全体严格三角形矩阵.

(h) 第二行元素为零的全体矩阵.

(i) 至少有一行的元素为零的全体矩阵.

5. 列举与F^n的所有平移构成的群同构的矩阵群.

6. (a) 设$\mathbf{Z}_2$是模2整数域，列出$L_2(\mathbf{Z}_2)$中所有矩阵.

(b) 构造群$L_2(\mathbf{Z}_2)$的乘法表.

*7. 当$\mathbf{Z}_p$是模p整数域时，全线性群$L_2(\mathbf{Z}_p)$的阶是多少？

8. 设G是所有形为$A=\begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$(其中$ad\neq 0$)的矩阵构成的群. 证明：对应$A\mapsto a$是一个同态.

9. 将3×3非奇异三角形矩阵构成的群同态地映射到2×2非奇异三角形矩阵构成的群.（提示：象习题8那样证明，但这里用分块矩阵.）

10. 证明：如果两个域F和K是同构的，那么群$L_n(F)$和$L_n(K)$也是同构的.

11. 设$n<m$，证明：$L_n(F)$与$L_m(F)$的子群同构.

12. (a) 证明：线性群$L_n(F)$的中心由标量矩阵$cI(c\neq 0)$组成.（提示：它们一定同每个矩阵$I+E_{ij}$可交换.）

(b) 证明：同每个仿射变换可交换的唯一的仿射变换是恒等变换.

*13. 设$L_n(F)$是全线性群，证明：两个仿射变换H_1和H_2落入$L_n(F)$的同一个右陪集中当且仅当$OH_1=OH_2$(O是原点).

14. 证明：商群$A_n(F)/T_n(F)$与$L_n(F)$同构，这里A_n表示仿射群，T_n表示平移群.

15. (a) 证明：满足$ad\neq bc$的全体一一变换$y=\dfrac{ax+b}{cx+d}$构成一个群(称为线性分式群).

(b) 证明：这个群与全线性群在模非零标量矩阵的子群之下的商群同构.

*(c) 对于大于 2×2 的矩阵，推广上述结果.

16. (a) 证明：所有形为$\begin{pmatrix} A & O \\ O & B \end{pmatrix}$(其中 A 为 $r\times r$ 矩阵，B 为 $s\times s$ 矩阵)的非奇异矩阵构成的集合与直积 $L_r(F)\times L_s(F)$ 同构.

(b) 当 $r=2, s=1$，用上述矩阵确定的 $\mathbf{R}^3$ 的线性变换的几何特征是什么?

§9.4 正交群与欧几里得群

在欧几里得几何中，长度这个概念起着极重要的作用. 因此我们在欧几里得矢量空间中寻求使得所有矢量 ξ 的长度 $|\xi|$ 保持不变的那些线性变换.

定义 *欧几里得矢量空间的线性变换 T，如果它保持每个矢量 ξ 的长度不变，即 $|\xi T|=|\xi|$，那么称 T 为正交变换.*

我们现在来确定欧几里得平面的所有正交变换 $Y=XA$. 因为 A 是正交变换，所以单位矢量 $(1, 0)$ 和 $(0, 1)$ 的变换式

$$(1, 0)\begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \end{pmatrix}=(a_1, a_2), \quad (0, 1)\begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \end{pmatrix}=(b_1, b_2) \qquad (17)$$

的长度为 1. 根据毕达哥拉斯长度公式，这就意味着

$$a_1^2+a_2^2=1, \quad b_1^2+b_2^2=1. \qquad (18)$$

此外，矢量 $(1, 1)$ 具有长度为 $\sqrt{2}$ 的变换式 (a_1+b_1, a_2+b_2)，所以 $(a_1+b_1)^2+(a_2+b_2)^2=2$，展开后再把 (18) 代入，我们得到

$$a_1b_1+a_2b_2=0. \qquad (18')$$

根据 (18)，存在一个角 θ 使得 $\cos\theta=a_1$，$\sin\theta=a_2$. 然后由 (18′) 有 $\operatorname{tg}\theta=\frac{a_2}{a_1}=-\frac{b_1}{b_2}$，因此根据 (18) 有 $b_2=\pm\cos\theta, b_1=\mp\sin\theta$. 正负号两种不同的选法恰好给出两个矩阵

$$\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \quad \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix} \tag{19}$$

根据§8.1公式(5)和(5′)，这两个矩阵分别表示转过θ角的旋转和关于与x轴成$\alpha=\frac{\theta}{2}$角的直线的反射．因此每个平面正交变换是旋转或是反射．

几何上，(19)式的左边一个正交变换的逆可通过把θ用$-\theta$代替而得到；因此这个逆是原来矩阵的转置．这个事实(不象三角公式)可以推广到$n\times n$正交矩阵．

定理 10 正交变换T具有性质：对每对矢量ξ,η，有

(i) T保持距离不变，即$|\xi-\eta|=|\xi T-\eta T|$．

(ii) T保持内积不变，即$(\xi,\eta)=(\xi T,\eta T)$．

(iii) T保持正交性，即由$\xi\perp\eta$可推出$\xi T\perp\eta T$．

(iv) T保持角的大小不变，即$\cos\angle(\xi,\eta)=\cos\angle(\xi T,\eta T)$．

证明 由于T是线性的，由定义得到(i)．因为$\xi\perp\eta$意味着$(\xi,\eta)=0$，而且角是通过内积来定义的(§7.9,(4))，所以性质(iii)和(iv)从性质(ii)直接推出．至于性质(ii)，从内积的“双线性”可证明$(\xi+\eta,\xi+\eta)=(\xi,\xi)+2(\xi,\eta)+(\eta,\eta)$．由这个方程可以利用“长度”(比如$|\xi|=(\xi,\xi)^{\frac{1}{2}}$)解出$(\xi,\eta)$，形为

$$2(\xi,\eta)=|\xi+\eta|^2-|\xi|^2-|\eta|^2. \tag{20}$$

由于正交变换T保持等式右边的长度不变，所以它也保持等式左边的内积不变．这就证明了(ii)． 证毕

反过来，如果已知变换T保持所有的内积不变，因为长度是通过内积来定义的，所以T一定保持长度不变，因此它是正交变换．

下面我们要问，什么样的矩阵对应于正交线性变换？这个问题至少对于标准正交基的情形是容易回答的．

定理 11 对于任意标准正交基，$n\times n$实矩阵A表示一个正交

线性变换当且仅当A的每个行矢量的长度为1，任意两个行矢量是正交的。

证明 根据定理10，任意正交变换T一定把已知基底$\varepsilon_1, \cdots, \varepsilon_n$变换为一组标准正交基$\alpha_1=\varepsilon_1T, \cdots, \alpha_n=\varepsilon_nT$。反过来，如果变换$T$具有这个性质，那么对任意矢量$\xi=x_1\varepsilon_1+\cdots+x_n\varepsilon_n$，有变换式$\xi T=x_1\alpha_1+\cdots+x_n\alpha_n$，由§7.11的定理22我们知道，长度是按普通公式给出

$$|\xi|=(x_1^2+\cdots+x_n^2)^{\frac{1}{2}}=|\xi T|,$$

因此T是正交的．根据下面的说明（参看§8.1）我们就完成了定理的证明：即A的第i行表示矢量$\alpha_i=\varepsilon_iT_A$关于原基底$\varepsilon_1, \cdots, \varepsilon_n$的坐标。

定理中所叙述的关于A的条件，写成坐标形式，等价于方程

$$\sum_{k=1}^{n} a_{ik}a_{ik}=1, \quad \text{对所有 } i,$$
$$\sum_{k=1}^{n} a_{ik}a_{jk}=0, \quad \text{当 } i\neq j. \tag{21}$$

对于2×2矩阵，结论(21)恰好就是在(18)和(18′)中已经建立的那些公式．如果我们用A_i表示矩阵A的第i行，用A_i^{τ}表示它的转置，A_i和A_j的内积是矩阵的乘积$A_iA_j^{\tau}$（见§8.5的(34)），那么条件(21)可以写成

$$A_iA_i^{\tau}=1, \qquad A_iA_j^{\tau}=0, \text{当 } i\neq j. \tag{21′}$$

在矩阵A与它的转置A^{τ}的乘积AA^{τ}中，按照行与列的相乘，(21′)式表明，第i行乘第j列是$A_iA_j^{\tau}=\delta_{ij}$，这里$\delta_{ij}$是单位矩阵$I=(\delta_{ij})$中的第$i$行第$j$列元素．单位矩阵的对角线元素$\delta_{ii}=1$，其他非对角线元素都是零．（记号$\delta_{ij}$称为克罗内克尔符号．）于是我们证明了

定理 12 一个 $n\times n$ 实矩阵表示正交变换当且仅当 $AA'=I$.

方程 $AA'=I$ 在任意域上都有意义，因此正交矩阵的概念可以定义得更一般些.

定义 任意域上的方阵 A 是正交的当且仅当 $AA'=I$.

这就意味着正交矩阵 A 的转置 A' 是 A 的右逆，因此根据 §8.6 定理 9, 每个正交矩阵 A 是非奇异的, 且满足 $A^{-1}=A'$. 因此 $A'A=I$. 这方程可写成 $A'(A')'=I$，因此 A' 是正交矩阵. 这就是说, 任意正交矩阵 A 的转置也是正交矩阵. 由此还可以推出: 矩阵 A 是正交的当且仅当 A 的每一列矢量的长度为 1, 任意两个列矢量正交:

$$\begin{aligned}&\sum_{k=1}^{n}a_{ki}a_{ki}=1, \quad \text{对所有 } i,\\&\sum_{k=1}^{n}a_{ki}a_{kj}=0, \quad \text{当 } i\neq j.\end{aligned}\tag{22}$$

所有 $n\times n$ 正交矩阵构成一个群. 这是显然的, 因为正交矩阵的逆 $A^{-1}=A'$ 是正交的，并且两个正交矩阵 A 和 B 的乘积是正交的: $(AB)'=B'A'=B^{-1}A^{-1}=(AB)^{-1}$. 这个群是全线性群 $L_n(F)$ 的子群，称它为*正交群* $O_n(F)$；当 $F=\mathbf{R}$ 时，$O_n(F)$ 与已知欧几里得空间的所有正交变换构成的群同构.

欧几里得矢量空间 E 的*刚体运动*指的是 E 中保持距离不变的非奇异变换 U，也就是说，对所有矢量 ξ,η,U 满足 $|\xi U-\eta U|=|\xi-\eta|$. E 的任意平移保持矢量差 $\xi-\eta$ 不变, 因此也保持它们的长度不变, 所以它是刚体运动. 因此，如果仿射变换 $\xi\longmapsto\xi T+\kappa$ 是刚性变换, 那么 $\xi\longmapsto(\eta-\kappa)=\xi T$ 也是刚性变换. 反过来, 如果 T 是刚性变换，那么 $\xi\longmapsto\eta=\xi T+\kappa$ 也是刚性变换. 而根据定理 10，线性变换是刚性的当且仅当它是正交变换，于是我们得出结论, *仿射变换* (11) *是刚体运动当且仅当 T 是正交变换*. 象在定理 7

的证明中那样，因为全体正交变换构成群，所以可以得出，全体刚性仿射变换构成仿射群的一个子群，它称为欧几里得群．这是欧几里得几何的基础①．

还有其他各种几何群．我们所熟悉的一个几何群是所有相似变换 T 构成的群，它是由使所有长度乘上数因子 $c_T>0$ 的线性变换 T 组成，所以 $|\xi T|=c_T|\xi|$．可以证明，这些相似变换实际上构成一个群，它包含正交群作为子群．“广义”相似群是由所有仿射变换 $\xi\longmapsto\xi T+\kappa$ 组成，其中 T 是相似变换．

习　题

1．检验下列矩阵的正交性．如果某个矩阵是正交的，求出它的逆矩阵：

(a) $\begin{pmatrix}\frac{1}{2} & \frac{\sqrt{3}}{2}\\ -\frac{\sqrt{3}}{2} & \frac{1}{2}\end{pmatrix}$　　(b) $\begin{pmatrix}\frac{1}{2} & \frac{\sqrt{3}}{2}\\ \frac{\sqrt{3}}{2} & \frac{1}{2}\end{pmatrix}$

(c) $\begin{pmatrix}0.6 & 0.8\\ 0.8 & -0.6\end{pmatrix}$

2．求出一个正交矩阵，它的第一行是矢量 $(5,12,0)$ 乘上一个标量．

3．证明：如果把正交矩阵的列置换，那么置换后的矩阵仍是正交矩阵．

4．证明：如果 A 和 B 都是正交矩阵，那么 $\begin{pmatrix}A & O\\ O & B\end{pmatrix}$ 和 $\begin{pmatrix}O & A\\ B & O\end{pmatrix}$ 也是正交矩阵．

5．把下面两个矩阵相乘，检验乘积矩阵的正交性：

$$\begin{pmatrix}\cos\phi & \sin\phi & 0\\ -\sin\phi & \cos\phi & 0\\ 0 & 0 & 1\end{pmatrix},\quad \begin{pmatrix}1 & 0 & 0\\ 0 & \cos\theta & \sin\theta\\ 0 & -\sin\theta & \cos\theta\end{pmatrix}.$$

6．证明：欧几里得群与矩阵群同构．

7．证明：全体平移构成欧几里得群的正规子群．

---

① 事实上，任意刚体运动必须是仿射变换，因此欧几里得群实际上是所有刚体运动的群．

8. 作为定理 10 中的性质(ii) 的另一个证明，用基本原理证明 $4(\xi, \eta) = |\xi+\eta|^2-|\xi-\eta|^2$.

9. 证明：一个仿射变换 H 同每个平移可交换当且仅当 H 本身是一个平移.

10. 证明：任意相似变换 S 只能按照一种方法写成形式 $S=cT$, 它是正标量 c 和正交变换 T 的乘积.

11. 对于标准正交基，给出矩阵 A 表示相似变换的充分必要条件.（参看定理 11 和定理 12.）

12. (a) 证明：全体相似变换构成群 S_n.

(b) 证明：O_n 是 S_n 的正规子群.

(c) 证明：商群 S_n/O_n 与全体正实数构成的乘法群同构.

13. $\mathbf{Z}_2$ 上的 3×3 矩阵中有多少个正交矩阵？$\mathbf{Z}_3$ 上呢？

14. (a) 证明：对应 $A\to\theta(A)=(A^{-1})^T$ 是全线性群 $L_n(F)$ 的一个自同构.

(b) 证明：对所有的 A, 有 $\theta^2(A)=A$.

(c) 对哪些矩阵 A, 有 $\theta(A)=A$？

§ 9.5 不变量与标准型

全线性群、仿射群、正交群和欧几里得群都是线性群的例子. 另一个例子是酉群 (§ 9.12). 在下面几节中，我们将看到使用这些群中适当的变换，多项式、二次型以及各种几何图形可以“简化”到什么程度. 这些化简类似于化简一般矩阵为行等价的简化梯矩阵，已经证明了它们的秩在所考虑的变换之下是一个不变量. 化简后的“标准型”和“不变量”的概念可以给出下面更一般的描述①.

设 G 是任意集合或“空间” S 上的变换群(§ 6.2). 我们称 S 上两个元素 x 和 y 在 G 之下是等价的(记作 xE_Gy) 当且仅当存在 G 的某一变换 T 把 x 变到 y. 那么 T^{-1} 把 y 变回 x, 所以 yE_Gx, 因此这个等价关系是对称的. 类似地，用群的其他性质我们可以证

① 读者注意，当你读完本章之后，再回来讨论这一般情形.

明，在任意G之下的等价性还满足自反律和传递律(即为等价关系)．对于S的子集合C，如果每个元素$x\in S$，在G之下与C中一个且只有一个元素c等价，那么C称为S在G之下的标准型集合；这个元素c就是x的标准型．对S的所有元素x定义的函数$F(x)$，它取值在另一个适当的集合上，比如说一个数集，如果它对于S中每个点x和G中每个变换T有$F(xT)=F(x)$，则称函数$F(x)$是G之下的不变量(有时称为不变式)；换句话说，F在所有等价元素上的值一定是相同的．一组不变量$F_1, \cdots, F_n$，如果可由$F_1(x)=F_1(y), \cdots, F_n(x)=F_n(y)$推出$x$和$y$等价，则称它们是$G$之下的全系不变量．

例如，设空间S是某一域上所有$n\times n$矩阵组成的集合M_n．对这些矩阵我们手边已经有三种不同等价关系，连同三种新的等价关系一起列在下面．后面三种新的等价关系将在以后几节(分别在§9.8，§9.10和§9.12)里讨论．

A行等价于B	$B=PA$，P为非奇异矩阵；
A等价于B	$B=PAQ$，P，Q为非奇异矩阵；
A相似于B	$B=PAP^{-1}$，P为非奇异矩阵；
A相合于B	$B=PAP'$，P为非奇异矩阵；
A正交等价于B	$B=PAP^{-1}$，P为正交矩阵；
A酉等价于B	$B=PAP^{-1}$，P为酉矩阵．

上面第一个等价关系读作："A行等价于B当且仅当存在非奇异矩阵P使得$B=PA$"．对其他各关系可给出类似的读法．

上述每个等价关系是由作用在M_n上的一个适当的群G所确定的等价关系E_G，而且自然产生于对矩阵的一种解释．

第一个行等价关系是在讨论把F^n的一个固定子空间表示成矩阵A的行空间中产生的．在这种情况中，矩阵P的全线性群通过$A\longmapsto PA$作用在A上，简化梯矩阵是这个群之下的标准型．A

的秩是这个群之下的不变量(数值的)，但它并没有给出全系不变量，这因为秩相同的两个矩阵 A 和 B 不一定是行等价的.

第二个等价关系(按 $B=PAQ$ 这种意义的等价，不会同一般的等价关系混淆)是在讨论一个矢量空间到另一个矢量空间的线性变换的各种矩阵表示中产生的(参看 § 9.2 习题 9). 这里，根据 § 8.9 的定理 18，秩是在群 $A\longmapsto PAQ$ 之下的全系不变量. 对角线元素是 1 或 0(1 都在 0 的前面)的所有对角矩阵组成的集合是标准型集合. 注意，我们同样还可以选取不同的标准型集合，比如说它们是同一类型的对角矩阵，只是对角线上的 0 在 1 的前面.

相似关系是在讨论一个矢量空间到自身的线性变换的各种矩阵表示中产生的，在这种情形，全线性群是通过 $A\longmapsto PAP^{-1}$ 作用到 A 上. 在相似变换之下，矩阵 A 的秩是一个不变量，因为两个相似矩阵是等价的，在等价之下，秩是不变量. 根据 §9.2，矩阵的所有特征值集合在相似变换之下也是不变量，但它不是全系不变量. 在相似变换之下，系统地列举出完全的标准型集合是矩阵论中一个重要的问题；对于复数域，给出了矩阵的若当标准型(见 § 10.10).

后面将出现的相合关系($B=PAP'$)，它是在讨论用(对称)矩阵表示二次型中产生的.

作为在群之下等价的另一个例子，我们考虑用所有平移 $y=x+k$ 构成的群来化简二次多项式 $f(x)=ax^2+bx+c$，其中 $a\neq 0$. 将 $x=y-k$ 代入，我们得到平移 $f(x)$ 的结果是

$$\begin{aligned} g(y) &= a(y-k)^2+b(y-k)+c \\ &= ay^2+(b-2ak)y+ak^2-bk+c. \end{aligned}$$

特别是，我们得到熟知的"配方"法——所得的新多项式没有一次项当且仅当 $k=\frac{b}{2a}$，这时，多项式是

$$g(y)=ay^2-\frac{d}{4a}, \text{其中 } d=b^2-4ac. \qquad (23)$$

于是 $f(x)$在平移群之下与一个且只与一个形为 ay^2+h 的多项式等价，因此没有一次项的二次多项式是这个群之下的标准型．另一方面，任意平移后的多项式同原多项式 $f(x)$具有相同的首项系数 a 和相同的判别式 $d=(b-2ak)^2-4a(ak^2-bk+c)$．因此$f(x)$的首项系数和判别式是这个群之下的不变量，它们组成全系不变量，因为标准型可以通过首项系数和判别式表示成(23)．

为了给出最后一个例子，回忆一下，全线性群 $L_n(F)$是矢量空间 F^n 的变换群．这个群的每个变换把 F^n 的子空间 S 变到另一个子空间．根据 §8.6 定理 10 的推论 2，任意子空间 S 的维数是全线性群之下的不变量．这一个不变量实际上是全线性群之下关于 F^n 的子空间的全系不变量(见下面习题 5)．

习　　题

1．求出所有的首项系数为 1 的二次多项式 x^2+bx+c 在平移群之下的标准型．

2．求出所有二次多项式 ax^2+bx+c (其中 $a\neq0$) 在仿射群 $y=hx+k$, $h\neq0$ 之下的标准型．

3．在习题 2 中，证明：$\frac{d}{a}=\frac{b^2}{a}-4c$ 是一个仿射不变量．

4．证明：在满足条件 $1+1\neq0$ 的任意域上，任意四次多项式在平移之下等价于一个没有三次项的多项式．

5．设 V 是 n 维矢量空间，证明：V 的子空间的有序对(S_1, S_2)在全线性群之下的全系不变量是由 S_1, S_2 的维数以及 S_1, S_2 的交的维数给出．

6．考虑在群 $x\mapsto rx$ (其中 $r\neq0$ 为有理数)之下的含有有理系数 a 的齐次二次函数 ax^2 的集合．证明：系数 a 为不同素数乘积 (无平方因子) 的二次函数的集合提供了上述二次函数集合的标准型．

7．设 $f(x)$是一个变量的任意多项式，证明：$f(x)$ 的次数和实根的个数是在仿射群之下的两个不变量．

8. 证明：对于n个变量的多项式，最高次项的系数是平移群之下的不变量.

9. 证明：一个实三次多项式在仿射群之下与一个且只与一个形为x^3+ax+b的多项式等价.

*10. 考虑二次函数x^2+bx+c，其系数b，c是模2整数域$\mathbf{Z}_2$中的元素，求出这种二次函数在平移群之下的标准型.

§9.6 线性型与双线性型

域F上的n个变量的线性型是形为

$$f(x_1, \cdots, x_n)=b_1x_1+\cdots+b_nx_n+c \tag{24}$$

的多项式，其中系数b_1，$\cdots$，b_n和c都在F中. 除了平凡情形之外，我们可假定某一个系数b_i不为零. 如果$c=0$，则这个线性型称为齐次线性型. 任意线性型(24)可以看作F^n的矢量$X=(x_1, \cdots, x_n)$的函数$f(X)$. 不同的线性型确定不同的函数，这因为函数$f(X)$通过公式

$$f(0, \cdots, 0)=c, f(1, 0, \cdots, 0)=b_1+c, \cdots, f(0, \cdots, 0, 1)=b_n+c$$

确定线性型(24)的系数.

对任意线性型，我们可以应用非奇异仿射变换

$$x_i=\sum_j a_{ij}y_j+k_i, \quad (a_{ij})\text{是非奇异的} \tag{25}$$

把它代入(24)中，产生新的线性型

$$g(y_1, \cdots, y_n)=\sum_j\left(\sum_i b_ia_{ij}\right)y_j+\left(\sum_i b_ik_i+c\right). \tag{26}$$

如果存在一个这样的仿射变换把f变到g，我们就说f和g在仿射群之下是等价线性型.

不难得到线性型的标准型. 首先，因为某个$b_j\neq0$，平移$x_j=y_j-\frac{c}{b_j}$，$x_i=y_i$(当$i\neq j$)将消去常数项. 置换$z_1=y_j$，$z_j=y_1$，$z_i=y_i$(当$i\neq1$和$i\neq j$)将给出象(24)那样一个新的线性型，其中b_1

$\neq 0$, 且 $c=0$. 如果这个线性型按变量 $x_1, \cdots, x_n$ 写出, 那么由方程

$$y_1=b_1x_1+\cdots+b_nx_n,\quad y_2=x_2, \cdots, y_n=x_n$$

给出的新的仿射变换是非奇异的, 它把满足 $c=0$ 的任意函数 f 变到等价的函数 $g(y_1, \cdots, y_n)=y_1$. 因此, *在仿射群之下, 所有非零线性型是等价的.*

现在考虑在欧几里得群(即(25)式中的$A=(a_{ij})$是正交矩阵)之下实线性型的等价性. $d=(b_1^2+\cdots+b_n^2)^{\frac{1}{2}}$称为线性型(24)的范数. 如上所述, 通过平移我们可以消去常数 c. 适当选取 d 使得 $\left(\frac{b_1}{d}, \cdots, \frac{b_n}{d}\right)$是具有单位长度的矢量, 因此存在一个正交矩阵$(h_{ij})$, 以上述矢量作为它的第一行. 那么变换 $y_i=\sum h_{ij}x_j$ 属于欧几里得群; 因为 $dy_1=b_1x_1+\cdots+b_nx_n$, 所以它把满足 $c=0$ 的线性型 f 变到线性型 $g=dy_1$.

这个线性型 dy_1 是线性型在欧几里得群之下的标准型. 为了证明这一点, 我们只须证明范数 d 是在欧几里得群下的不变量. f 的范数 d 正好是系数矢量 $\beta=(b_1, \cdots, b_n)$的长度, (26)式表明, 变换后的线性型中的系数矢量是原来的系数矢量在正交矩阵(a_{ij})作用之下的变换式 βA; 因此范数是不变量. 于是我们证明了

定理 13 *在欧几里得群之下, 每个线性型(24)与一个且只与一个标准型dy等价, 其中 d 是满足 $d=(b_1^2+\cdots+b_n^2)^{\frac{1}{2}}$的正数, 它是这个群之下的不变量.*

关于两组变量 $x_1, \cdots, x_m$ 和 $y_1, \cdots, y_n$ 的(齐次)双线性型是形为

$$b(x_1, \cdots, x_m, y_1, \cdots, y_n)=\sum_{i=1}^{m}\sum_{j=1}^{n}x_ia_{ij}y_j \tag{27}$$

的多项式, 它是通过系数矩阵 $A=(a_{ij})$来确定的. 这个双线性型

可以利用矢量 $X=(x_1, \cdots, x_m)$ 和 $Y=(y_1, \cdots, y_n)$ 写成矩阵乘积

$$b(X, Y)=XAY'. \tag{28}$$

作为 X 和 Y 的函数，这个函数分别对每个自变量是线性的.

更一般地，设 V 和 W 是同一个域 F 上的维数分别为 m 和 n 的有限维矢量空间，且设 $B(\xi, \eta)$ 是对于自变量 $\xi\in V$ 和 $\eta\in W$ 定义的，取值在 F 上的任意函数，它按下述意义是双线性的：对于 a_1 和 $a_2\in F$，有

$$B(a_1\xi_1+a_2\xi_2, \eta)=a_1B(\xi_1, \eta)+a_2B(\xi_2, \eta),$$
$$\xi_1, \xi_2\in V, \quad \eta\in W; \tag{29}$$
$$B(\xi, a_1\eta_1+a_2\eta_2)=B(\xi, \eta_1)a_1+B(\xi, \eta_2)a_2,$$
$$\xi\in V, \quad \eta_1, \eta_2\in W. \tag{29'}$$

选取 V 的基底 $\alpha_1, \cdots, \alpha_m$ 和 W 的基底 $\beta_1, \cdots, \beta_n$，且设 a_{ij} 是按 $a_{ij}=B(\alpha_i, \beta_j)$ 定义的标量. 那么对 V 和 W 中任意矢量 ξ 和 η，按照基底表示，我们有

$$B(\xi, \eta)=B(x_1\alpha_1+\cdots+x_m\alpha_m, y_1\beta_1+\cdots+y_n\beta_n).$$

因此根据(29)和(29')得

$$B(\xi, \eta)=\sum_{i,j} x_iB(\alpha_i, \beta_j)y_j=\sum_{i,j}x_ia_{ij}y_j.$$

换句话说，V 和 W 上的任意双线性函数 B 对于给定的一组基底有象 (27) 那样的唯一表达式. 按照 § 8.5 的记号，另一等价的说法是，双线性型恰好是 m 维行矢量 X, $m\times n$ 矩阵 B 和 n 维列矢量 Y 的乘积 XBY.

这两个空间的基底变换对应着各组变量的非奇异变换 $X=X^*P$ 和 $Y=Y^*Q$. 在这些变换下，可用新的双线性型 $X^*(PAQ')Y^*$ 代替(28)式，其中 PAQ' 是一个新的矩阵. 因为任意非奇异矩阵可以写成一个非奇异矩阵的转置 Q'，所以我们看出，两个双线性型(在基底变换之下)是等价的当且仅当它们的矩阵是等价

的. 因此, 根据§8.9中关于矩阵等价性的定理18, 可知任意双线性型与一个且只与一个标准型

$$x_1y_1+\cdots+x_ry_r$$

等价. 这里整数 r 是双线性型的矩阵的秩, 它是一个(全系)不变量.

习　题

1. 求出齐次实线性函数在相似群之下的标准型.

2. 分别在下面两种情形下求齐次实线性函数的标准型:

(a) 在对角线变换群之下, 即 $y_1=d_1x_1, \cdots, y_n=d_nx_n$.

(b) 在单项变换群之下, 即 $Y=XM$, 其中 M 为单项矩阵.

3. 证明: 秩为 r 的任意双线性型可以表示成

$$\sum_{i=1}^{r}(b_{i1}x_1+\cdots+b_{in}x_n)(c_{i1}y_1+\cdots+c_{in}y_n),$$

即表示成 r 个线性型的乘积之和.

4. 求出两组新变量 x^*, y^*, z^* 和 u^*, v^*, w^*, 把下面双线性函数化为标准型:

$$xu+xv+xw+yu+yv+yw+zu+zv+zw.$$

§9.7 二　次　型

下面四节专门研究二次型在各种变换群之下的标准型. 这类问题中最简单的是产生于对平面有心二次曲线(具有"斜"轴的椭圆或双曲线)的研究. 这样的二次曲线满足方程 $Ax^2+Bxy+Cy^2=1$, 其中左边是"二次型". 这样的二次型(全体变量的二次齐式)产生于很多其他情形: 例如, 空间的二次曲面方程, 二次曲线在齐次坐标下的射影方程, 矢量长度的平方公式 $|X|^2=x_1^2+x_2^2+\cdots+x_n^2$, 具有三个速度分量 u, v, w 的空间运动物体的动能公式 $\frac{m}{2}(u^2+v^2+w^2)$, 微分几何中在球面坐标下空间的弧长 ds 的公式 $ds^2=$

$dr^2+r^2d\phi^2+r^2\sin\phi d\theta^2$.

这样的二次型可以用矩阵表示. 例如, 为了得到二次型 $5x^2+6xy+2y^2$ 的矩阵表示, 首先调整二次型使得 xy 和 yx 的系数相等, 写成 $5x^2+3xy+3yx+2y^2$. 这个表达式可以写成矩阵乘积

$$(x, y)\begin{pmatrix}5 & 3\\ 3 & 2\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}=(x, y)\begin{pmatrix}5x+3y\\ 3x+2y\end{pmatrix}$$
$$=5x^2+6xy+2y^2.$$

这里给出的 2×2 系数矩阵是对称的. 由于对称性, 这个矩阵等于它的转置.

一般地, 如果方阵 A 等于它的转置: $A^{\tau}=A$, 则称 A 是对称的; 换句话说, (a_{ij}) 是对称的当且仅当对所有的 i, j, 有 $a_{ij}=a_{ji}$. 类似地, 如果矩阵 C 满足 $C^{\tau}=-C$, 则称 C 是斜对称的. 为把矩阵 B 分成对称部分和斜对称部分, 我们可以把 B 写成

$$B=\frac{B+B^{\tau}}{2}+\frac{B-B^{\tau}}{2}=S+K, \tag{30}$$

其中 $S=\frac{B+B^{\tau}}{2}$, $K=\frac{B-B^{\tau}}{2}$. 根据转置的运算法则, 有 $(B\pm B^{\tau})^{\tau}=B^{\tau}\pm(B^{\tau})^{\tau}=B^{\tau}\pm B$, 所以 S 是对称的, K 是斜对称的. 不可能再有别的分解使得 $B=S_1+K_1$, 其中 S_1 是对称的, K_1 是斜对称的. 这是因为任何这样的分解将给出 $B^{\tau}=S_1^{\tau}+K_1^{\tau}=S_1-K_1$, 于是 $B+B^{\tau}=2S_1$, $B-B^{\tau}=2K_1$, 所以 $S_1=S$, $K_1=K$. 公式(30)可以应用到任意满足条件 $2=1+1\neq 0$ 的域上, 但对于域 $\mathbf{Z}_2$ 上的矩阵, 公式(30)没有意义, 因为在 $\mathbf{Z}_2$ 中 $1+1=0$. 总之, 任意矩阵可以唯一地表示成对称矩阵与斜对称矩阵之和, 只要假定 $1+1\neq 0$.

n 个变量 $x_1, \cdots, x_n$ 的齐次二次型是由多项式

$$\sum_i\sum_j x_i b_{ij} x_j$$

来定义的, 其中每一项的次数都是 2. 这个二次型可以写成矩阵

的乘积 XBX'. 如果系数矩阵 B 是斜对称的，那么 $b_{ij}=-b_{ji}$，因此这个二次型等于零. 一般地，根据(30)式，把 B 写成 $B=S+K$，二次型变成

$$XBX'=X(S+K)X'=XSX'+XKX'=XSX',$$

(K 是斜对称矩阵).

因此，当 $1+1\neq 0$，则任意二次型可以唯一地表示成(用 A 记 S)

$$\sum_{i=1}^{n}\sum_{j=1}^{n}x_i a_{ij} x_j = XAX', \quad A=(a_{ij})\text{是对称矩阵}. \tag{31}$$

如果矢量 ξ 有坐标 $X=(x_1,\cdots,x_n)$，那么每个二次型确定一个矢量 ξ 的二次函数 $Q(\xi)=XAX'$. 由空间的基底变换给出新坐标 X^*，它与老坐标的关系是通过方程 $X=X^*P$(其中 P 是非奇异矩阵)给出的. 按照 ξ 的新坐标来表示，二次型变成

$$Q(\xi)=XAX'=(X^*P)A(X^*P)'=X^*(PAP')X^{*\prime}$$

这是含有新矩阵 PAP' 的另一个二次型. 这个新矩阵同 A 一样是对称的，因为 $(PAP')'=(P')'A'P'=PAP'$.

定理 14 通过坐标变换，含有矩阵 A 的二次型变为含有矩阵 PAP' 的二次型，这里 P 是非奇异的.

对称矩阵 A 和 B，如果满足关系 $B=PAP'$(其中 P 是非奇异的)，则称 A 和 B 是相合的.

再重复一下，定理 14 断言，齐次的二次型在全线性群(这个群由关于变量的非奇异的线性齐次变换构成)之下化为标准型的问题，等价于求对称矩阵 A 在变换群 $A\longmapsto PAP'$ 之下的标准型问题.

习　题

1. 证明：$A'A$ 和 AA' 总是对称的.
2. 证明：如果 A 是斜对称的，那么 A^2 是对称的.

3. 把 § 8.3 中习题 1 的每个矩阵表示成 $S+K$ 的形式。

4. 求出与下列各二次型相联系的对称矩阵:

(a) $2x^2+3xy+6y^2$,

(b) $8xy+4y^2$,

(c) $x^2+2xy+4xz+3y^2+yz+7z^2$,

(d) $4xy$,

(e) $x^2+4xy+4y^2+2xz+z^2+4yz$.

5. (a) 证明: 如果 S 是对称的, A 是正交的, 那么 $A^{-1}SA$ 是对称的.

(b) 证明: 如果 K 是斜对称的, A 是正交的, 那么 $A^{-1}KA$ 是斜对称的.

6. 讨论矩阵 $AB-BA$ 在下列各情况下的对称性:

(a) A 和 B 都是对称的.

(b) A 和 B 都是斜对称的.

(c) A 是对称的, 而 B 是斜对称的.

7. 证明: 如果 A 和 B 是对称的, 那么 AB 是对称的当且仅当 $AB=BA$.

8. (a) 证明: 域 $\mathbf{Z}_2$(模 2 整数域)上每个斜对称矩阵是对称的.

(b) 举出 $\mathbf{Z}_2$ 上的矩阵, 它不能表示成和 $S+K$(见(30)式).

9. 设 D 是无重复元素的对角矩阵, 证明: $AD=DA$ 当且仅当 A 也是对角矩阵.

10. 设 $Q(\xi)$ 是二次函数, 证明

$$Q(\alpha+\beta+\gamma)-Q(\alpha+\beta)-Q(\beta+\gamma)-Q(\gamma+\alpha)+Q(\alpha)+Q(\beta)+Q(\gamma)=0.$$

11. 双线性型 $B(\xi,\eta)$(ξ 和 $\eta\in V$)如果满足 $B(\xi,\eta)=B(\eta,\xi)$, 则称它是对称的. 证明: 如果 B 是对称双线性型, 那么 $Q(\xi)=B(\xi,\xi)$ 是满足关系式

$$2B(\xi,\eta)=Q(\xi+\eta)-Q(\xi)-Q(\eta)$$

的二次型.

12. 证明: $n\times n$ 实矩阵 A 是对称的当且仅当与它相联系的 n 维欧几里得空间的线性变换 $T=T_A$ 对任意两个矢量 ξ,η 满足关系 $(\xi T,\eta)=(\xi,\eta T)$.

*13. 证明: 如果实矩阵 S 是斜对称的, 并且 $I+S$ 是非奇异的, 那么 $(I-S)(I+S)^{-1}$ 是正交矩阵.

§9.8 全线性群之下的二次型

大家熟悉的"配方"法可以用作为化简二次型(通过线性变换)的方法. 对于两个变量的二次型,由这个方法可以得到

$$
\begin{aligned}
& ax^2+2bxy+cy^2 \\
& \quad =a\left[x^2+2\frac{b}{a}xy+\frac{b^2}{a^2}y^2\right]+\left[c-\frac{b^2}{a}\right]y^2 \\
& \quad =a\left[x+\frac{b}{a}y\right]^2+\left[c-\frac{b^2}{a}\right]y^2.
\end{aligned}
$$

括号中的项给出新的变量 $x'=x+\frac{b}{a}y$, $y'=y$. 在这个变量线性变换下, 二次型变成 $ax'^2+\left[c-\frac{b^2}{a}\right]y'^2$, 交叉项被去掉了.

这一论证要求 $a\neq 0$. 如果 $a=0$, 而 $c\neq 0$, 则可用类似的变换来化简. 最后, 如果 $a=c=0$, 那么原来的二次型是 $2bxy$, 对应的方程 $2bxy=1$ 表示一个等轴双曲线. 在这种情况下, 变换 $x=x'+y'$, $y=x'-y'$ 将把二次型化为

$$2b(x'+y')(x'-y')=2b(x'^2-y'^2).$$

这个表达式也只包含着平方项. (提示: 这里使用的变换与关于双曲线轴的旋转有什么关系?)

类似的"配方"法可以应用到多于两个变量的二次型上.

引理 *通过非奇异线性变换, 任意不恒为零的二次型 $\sum x_i a_{ij} x_j$ 可以化为首项系数 $a_{11}\neq 0$ 的二次型, 只要假定 $1+1\neq 0$.*

证明 根据假设, 至少有一个系数 $a_{ij}\neq 0$. 如果是对角线元素 $a_{ii}\neq 0$, 那么我们通过 x_1 与 x_i 变量对换(这是一个非奇异变换, 因为它的矩阵是置换矩阵), 可以得到新的系数 $a'_{11}\neq 0$. 如果所有对角线元素 a_{ii} 都是零, 但是存在下标 $i\neq j$ 满足 $a_{ij}\neq 0$. 通过置换变量, 我们可以使 $a_{12}\neq 0$; 由矩阵的对称性有 $a_{12}=a_{21}$. 那么已

知的二次型就是 $a_{12}x_1x_2+a_{21}x_2x_1=2a_{12}x_1x_2$ 再加上含有其他变量的一些项．恰好同等轴双曲线的情形一样，通过变换

$$x_1=y_1-y_2,\quad x_2=y_1+y_2,\quad x_3=y_3,\ \cdots,\quad x_n=y_n,$$

这个二次型就可化成首项系数 $2a_{12}\neq 0$ 的形式 $2a_{12}(y_1^2-y_2^2)$．上述变换是非奇异的，由消去法我们容易证明它有逆变换

$$y_1=\frac{x_1+x_2}{2},\quad y_2=\frac{x_2-x_1}{2},\quad y_3=x_3,\ \cdots,\quad y_n=x_n.$$

问：这个论证中什么地方用到假设 $1+1\neq 0$？

现在我们对任意二次型"配方"．根据引理，我们可使 $a_{11}\neq 0$，所以二次型可以写成 $a_{11}(\sum x_ib_{ij}x_j)$，这里 $b_{ij}=\dfrac{a_{ij}}{a_{11}}$，$b_{11}=1$．由于矩阵的对称性，含有 x_1 的项为

$$x_1{}^2+2\sum_{j=2}^{n}b_{1j}x_1x_j=\left(x_1+\sum_{j=2}^{n}b_{1j}x_j\right)^2-\left(\sum_{j=2}^{n}b_{1j}x_j\right)^2.$$

这个"完全平方"的构成暗示着一个变换

$$y_1=x_1+\sum_{j=2}^{n}b_{1j}x_j,\quad y_2=x_2,\ \cdots,\ y_n=x_n.$$

那么 y_1 只出现在 y_1^2 项中．原来的二次型现在就变为 $a_{11}y_1^2+\sum y_jc_{jk}y_k$，这里的下标 j 和 k 是从 2 跑到 n．剩下的部分是 $n-1$ 个变量 $y_2,\cdots,y_n$ 的二次型，对这个二次型使用同样的方法．这个方法可以重复进行下去(归纳论证)直到剩下的二次型的新系数全都是零为止．因此我们有

定理 15 通过变量的非奇异线性变换，满足条件 $1+1\neq 0$ 的任意域上的二次型可以化为对角二次型

$$d_1y_1^2+d_2y_2^2+\cdots+d_ry_r^2,\quad 每个\ d_i\neq 0. \tag{32}$$

非零对角线元素的个数 r 是一个不变量．

这个数 r 称为已知二次型 XAX' 的秩．因为 r 是化简后的二

次型(32)的对角矩阵D的秩，所以它的不变性是显然的．这个秩必等于原来二次型的矩阵A的秩，因为根据定理 14，我们的变换把A化为$D=PAP^T$，而我们已经知道(§8.9 定理 19)，秩在更一般的变换$A\mapsto PAQ$之下是一个不变量．

如果n个变量的二次型XAX^T的秩是n，那么称它是非奇异的，因为这意味着矩阵A是非奇异的．

在对角二次型(32)中，秩r是不变量，而系数并不是不变量，因为用不同的方法化简二次型可以产生不同的系数组$d_1, \cdots, d_r$．下一节我们将得到实数域这个特殊情形下的全系不变量．

习　题

1. 在有理数域上，把§9.7 习题 4 中的每个二次型化为对角型．

2. 在模 5 整数域$\mathbf{Z}_5$上，把$2x^2+xy+3y^2$化成对角型．

3. 证明：在域$\mathbf{Z}_5$上，每个二次型可以通过线性变换化成形式$\sum d_i y_i^2$，其中每个系数$d_i=0, 1$或 2．

4. 证明：在有理数域上，二次型$x_1^2+x_2^2$可以变换成两种不同的对角型：$9y_1^2+4y_2^2$和$2z_1^2+8z_2^2$．

5. 当矩阵A是下列情形时，求出矩阵P使得PAP^T是对角矩阵．

(a) $A=\begin{pmatrix} 3 & 1 \\ 1 & 0 \end{pmatrix}$　　(b) $A=\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$

(c) $A=\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 2 \\ 0 & 2 & 0 \end{pmatrix}$

6. 求出所有的把实二次型$x_1^2+\cdots+x_n^2$化为$y_1^2+\cdots+y_n^2$的线性变换．

7. 严格证明：二次型xy在群$L_2(\mathbf{Z}_2)$之下不等价于对角型．

§9.9　全线性群之下的实二次型

在解析几何中，二次曲线和二次曲面都是用实二次多项式函数来描述．在实数域上，对角型(32)的每一项可以通过变量代换

$y_i' = |d_i|^{\frac{1}{2}} y_i$ 进一步化简，使得项 $d_i y_i^2$ 变成 $\pm y_i'^2$. 对所有变量同时实行代换，就把二次型化为形式 $\sum \pm y_i^2$. 在这个和中，可以置换变量，使得正的平方项都在前面. 这就证明了

定理 16 实数域上任意二次型可以通过变量的非奇异线性变换化成如下形式：

$$z_1^2 + \cdots + z_p^2 - z_{p+1}^2 - \cdots - z_r^2. \tag{33}$$

定理 17 出现在简化形式(33)中正平方项的个数 p 是给定二次型 Q 的不变量. 也就是说，p 只依赖于二次型而不依赖于化简的方法(西耳维斯特(Sylvester)惯性律).

证明 假设存在另外的简化形式

$$y_1^2 + \cdots + y_q^2 - y_{q+1}^2 - \cdots - y_r^2, \tag{34}$$

它含有 q 个正项. 因为这两个简化形式是从同一个 Q 通过非奇异变换得到的，所以存在一个非奇异变换把(33)化成(34). 我们可以把这个变换方程看成坐标变换(图象固定坐标移动的变换)，那么(33)和(34)表示固定矢量 ξ 的同一个二次函数 $Q(\xi)$，而 ξ 相对于一组基底的坐标是 $z_1, \cdots, z_r$，相对于另一组基底的坐标是 $y_1, \cdots, y_r$.

假定 $q<p$. 如果公式(33)中 $z_{p+1}=\cdots=z_r=0$，则 $Q(\xi)\geqslant 0$. 满足这 $r-p$ 个方程的全体矢量 ξ 构成 $n-(r-p)$ 维子空间 S_1 (在这个子空间中存在 $n-(r-p)$ 个坐标 $z_1, \cdots, z_p, z_{r+1}, \cdots, z_n$). 类似地，(34)式中，每个 $\xi \neq 0$ 如果坐标满足 $y_1=\cdots=y_q=y_{r+1}=\cdots=y_n=0$，则使得 $Q(\xi)<0$. 这些条件确定一个 $r-q$ 维子空间 S_2. S_1 和 S_2 两个子空间的维数之和是

$$n-(r-p)+(r-q)=n+(p-q)>n$$

因此 S_1 和 S_2 具有公共的非零矢量 ξ，因为由 § 7.8 定理 17 可知交 $S_1 Q S_2$ 的维数是正的. 对于这个公共矢量 ξ，由(33)有 $Q(\xi)\geqslant$

0，而由(34)有$Q(\xi)<0$，显然得出矛盾．如果假定$q>p$，则将导出类似的矛盾，所以$q=p$，定理证完．

这个结果表明，任意实二次型可以通过线性变换化为一种且只化为一种(33)那样的二次型．所以这种类型的表达式$\sum \pm z^2$是实二次型在全线性群之下的标准型．这个标准型本身由所谓符号差$\{+,\cdots,+,-,\cdots,-\}$唯一确定，这个符号差是由$p$个正号，$r-p$个负号组成，其中$r$是二次型的秩．这个符号差是通过$r$和$s=p-(r-p)=2p-r$确定的($s$是正号个数减去负号个数)．有时称这个整数$s$为二次型的符号差．$r$和$s$一起构成全系(数值)不变量，这是根据两个二次型等价当且仅当它们可化成同一个标准型(33)．

定理 18 两个实二次型在全线性群之下等价当且仅当它们有相同的秩和符号差．

n个变量的实二次型$Q=XAX'$，如果当$X\neq 0$时可推出$Q>0$，那么称Q是正定二次型；在同样条件下实对称矩阵A称为正定矩阵．如果我们考虑标准型(33)，显然，Q是正定二次型当且仅当标准型是$z_1^2+\cdots+z_n^2$．这是因为，n个平方和除了所有项都是零而外总是正的，而且当取X是第n个单位矢量e_n时，(33)中的$XAX'\leqslant 0$，除非$p=n$．于是，我们证明了

定理 19 实二次型是正定的当且仅当它的标准型是$z_1^2+\cdots+z_n^2$．

根据定理 14，这意味着$A=PIP'$，它给出下面进一步的结果．

定理 20 实对称矩阵A是正定的当且仅当存在一个实非奇异矩阵P使得$A=PP'$．

二次型XAX'确定了在n维矢量空间中的一条轨迹，它是由满足$XAX'=1$的所有点X组成．标准型(33)意味着，通过适当的非奇异线性变换可把轨迹化为具有方程

$$z_1^2+\cdots+z_p^2-z_{p+1}^2-\cdots-z_r^2=1$$

的曲线. 例如,在平面上,秩为 2 的简化了的方程是

$$x^2+y^2=1,\quad x^2-y^2=1,\quad -x^2-y^2=1.$$

它们分别表示圆,等轴双曲线,或根本没有轨迹. 秩为 0 的唯一的二次型是 $0=1$; 秩为 1 的二次型是 $x^2=1$(它表示两条直线 $x=\pm1$)或者 $-x^2=1$(没有轨迹). 在 § 8.8 中我们证明了(定理 15 的推论 2),平面上任意非奇异线性变换可以表示为切变换、压缩(或伸长)与反射的乘积. 因此方程 $ax^2+bxy+cy^2=1$ 表示的任意有心二次曲线可以相继通过切变换、压缩与反射化为我们在上面列举出的几种形式之一. 在几何上,这些结果是合理的:椭圆可以沿着一根轴压缩成圆;但是显然找不到一系列线性变换可以把圆 $x^2+y^2=1$ 化成等轴双曲线 $x^2-y^2=1$. 这就是平面情形符号差的不变性的几何意义.

在研究两个变量的函数的极大值和极小值时, 符号差是有用的. 设 $z=f(x,y)$ 是一个光滑函数, 它的一阶偏导数 f_x 和 f_y 在 $x=x_0$, $y=y_0$ 处都为零,因此 z 按 $h=x-x_0$, $k=y-y_0$ 的幂的泰勒级数展开式中,没有一次项. 这个展开式是

$$f(x_0+h,y_0+k)=f(x_0,y_0)+\frac{1}{2}[ah^2+2bhk+ck^2]+\cdots,$$

其中系数 a, b, c 是偏导数

$$a=f_{xx}(x_0,y_0),\quad b=f_{xy}(x_0,y_0),$$
$$c=f_{yy}(x_0,y_0).$$

当 h 和 k 的值很小时,起支配作用的是方括号中的那项,它是变量 h 和 k 的实系数二次型. 如果这个二次型的秩是 2,那么它可以利用变换后的变量 h' 和 k' 表示成 $\pm h'^2\pm k'^2$. 如果两项的符号都是加号, 那么(x_0,y_0)点附近的函数值 $f(x_0+h,y_0+k)$总是超过 $f(x_0,y_0)$,所以 z 有相对极小值. 如果两项的符号都是减号, 那么

z 有相对极大值．如果一项符号是加号另一项符号是减号，则二次型可以取正值也可以取负值，所以(x_0, y_0)既不是极大点也不是极小点，而是鞍点（象马鞍，或者象两个山峰之间的隘口，高度 z 沿一个方向是增加的，而沿另一个方向是减少的）．因此，极大点、极小点和鞍点是根据二次型的符号差来区分的．三个变量或更多变量的函数的临界点有类似的结果．

习　　题

1．证明：实二次函数 $ax^2+bxy+cy^2$ 是正定的当且仅当 $a>0$ 且 $4ac-b^2>0$．

2．证明：正定对称矩阵的主对角线上的元素都是正的．

3．把下列各实二次型化为定理 16 所述的标准型，并求出每个二次型的秩和符号差：

(a) $9x_1^2+12x_1x_2+79x_2^2$，

(b) $2x_1^2-12x_1x_2+18x_2^2$，

(c) $-2x_1^2-4x_1x_2+22x_2^2+12x_2x_3+6x_3x_1-x_3^2$．

4．描述三维空间中实二次型的各种可能标准型的几何轨迹．

5．证明：复系数齐次二次型在全复线性群之下总是与平方和 $z_1^2+\cdots+z_r^2$ 等价．

6．证明：n 个变量的复系数的两个二次型在全线性群之下等价当且仅当它们有相同的秩．

7．证明：双线性函数 XAY^T 是"内积"当且仅当 A 是对称的并且是正定的．

8．如果二次型的秩等于符号差，那么称这个二次型是半正定的，对这种二次型叙述并证明类似于定理 19 的命题．

9．对半正定二次型，叙述并证明类似于定理 20 的命题．

10．(a) 列出四个变量非奇异二次型的所有类型．

(b) 至少对两个二次型，几何地描述一下它们在 $\mathbf{R}^4$ 中相应的轨迹．

§9.10 正交群之下的二次型

实二次型在正交变换之下可以化简成什么样呢？一个正交变换 $Y=XP$ 把 XAX' 变换成 $Y(P^{-1}A(P^{-1})')Y'$. 因为 P 是正交矩阵，所以新矩阵可以写成① $P^{-1}A(P^{-1})'=P^{-1}AP$.

在平面上，椭圆在正交变换(旋转和反射)之下决不能产生圆；我们至多可以使椭圆的轴旋转到标准位置上. 长轴可以看作最长的直径. 为重述这个最大值性质，考虑任意实二次函数 $Q(\xi)=ax^2+cy^2$，其中 $a\leqslant c$，并且没有 xy 项. 那么 $Q(\xi)\leqslant cx^2+cy^2=c(x^2+y^2)$，这就意味着，在单位圆 $x^2+y^2=1$ 的所有点上 Q 的最大值是 c，并且在 $x=0, y=1$ 点上取这个最大值. 反过来，下面的引理保证 Q 中没有 xy 项.

引理 如果实二次函数 $Q=ax^2+2bxy+cy^2$ 在单位圆 $x^2+y=1$ 的所有点中有一个最大值，并且在点 $x=0, y=1$ 上取得，那么 $b=0$.

证明 把 Q 看作一个变量 x 的(双值)函数，这里 y 同 x 的关系隐含在等式 $x^2+y^2=1$ 中. 两边求微商，我们得到 $2x+2y\dfrac{dy}{dx}=0$，所以导数 $y'=\dfrac{dy}{dx}=-\dfrac{x}{y}$. Q 的导数是

$$Q'=(ax^2+2bxy+cy^2)'=2ax+2by+2bxy'+2cyy'.$$

置 $x=0$, $y=1$，求出 y' 的值后一起代入上式，我们得到 $Q'=2b$. 但是 Q 在 $x=0, y=1$ 点处达到最大值，所以这个导数一定为零，因此 $2b=0$.

证毕

现在回到 n 个变量的二次型. 在 n 维空间中，单位超球面 $\sum x_i^2=1$ 是封闭的有界集合，它上面的点都是长度为 1 的矢量.

---

① 两个对称矩阵 A 和 $P^{-1}AP$(P 是正交矩阵) 有时称为是正交相合的.

在这个超球面上，实二次型 $Q(\xi)=\sum_{i,j}x_ia_{ij}x_j$ 所取的值有一个上界 $\sum_{i,j}|a_{ij}|$. 因为 $Q(\xi)$是 ξ 的连续函数，所以 $Q(\xi)$在 S 上有最大值①λ_1. 换句话说，在所有单位长度的矢量 ξ 中间，存在一个矢量 ξ_0，在 ξ_0 上，$Q(\xi)$取它的最大值 λ_1. 因为 ξ_0 的长度为 1，所以我们可以选取 $\alpha_1=\xi_0$ 作为新的标准正交基 α_1, …, α_n 的第一个矢量（§ 7.11 的定理 21）. 设 ξ 对于这组基底的新坐标是 $y_1, \cdots, y_n$，那么二次型按照新坐标可表示为 $Q(\xi)=\sum y_ib_{ij}y_j$，其中系数 b_{ij} 组成一个新矩阵. Q 的最大值 λ_1 通过坐标为(1, 0, …, 0)的矢量 α_1 给出，所以代入 α_1 可知最大值 λ_1 就等于 b_{11}. 如果我们进一步限制变量，除了 y_1 和 y_i两个变量之外其他都是零，那么这个最大值仍然保持不变. 因此 $y_1=1$, $y_i=0$ 是二次型 $b_{11}y_1^2+2b_{1i}y_1y_i+b_{ii}y_i^2$ 在条件 $y_1^2+y_i^2=1$ 之下的极大点. 那么引理(用 y_i 代替 x)断言，交叉乘积的系数 b_{1i} 是零，这种论证应用于 $i=2, 3, \cdots, n$ 每一种情形. 因此，Q 按照这些坐标 $y_1, \cdots, y_n$ 来表示时，去掉了所有 y_i 与 y_1 交叉乘积项，于是变成

$$Q(\xi)=\lambda_1y_1^2+\sum_{i=2}^{n}\sum_{j=2}^{n}y_ib_{ij}y_j,\quad B=(b_{ij})=B^*. \tag{35}$$

第一个系数 λ_1 不是矢量而是标量(是在球面$|\xi|=1$上 $Q(\xi)$的最大值).

在(35)式中差 $Q^*(\xi)=Q(\xi)-\lambda_1y_1^2$ 是 $n-1$ 个变量 $y_2, \cdots, y_n$ 的二次型. 这些变量是由 $n-1$ 个新基矢量 α_2, …, α_n张成的空间 S_{n-1} 中的坐标. 在这个空间(它是第一个基矢量 ξ_0 的正交补)中，我们可以重复应用同样的方法，选择新的标准正交基使得 $Q^*(\xi)$

① 同微积分学中一样，这里我们假定下述事实成立：在有界闭集上的连续函数在这个集合上有一个最大值.

在$|\xi|=1$上取最大值；这就从二次型中分出另外一个对角线元素．最后我们求出一组主轴基底，对于这组基底有

$$Q(\xi)=\lambda_1 z_1^2+\lambda_2 z_2^2+\cdots+\lambda_n z_n^2, \quad \lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_n. \tag{36}$$

这里$z_1, \cdots, z_n$是ξ相对于$\alpha_1, \beta_2, \gamma_3, \cdots$这组基底的坐标，而$\alpha_1, \beta_2, \gamma_3, \cdots$是在逐次满足最大要求时选取的．第一个矢量$\alpha_1$使$Q(\xi)$达到最大值$\lambda_1$，只在条件$|\xi|=1$之下．第二个基矢量$\beta_2$选为在这个空间中与$\alpha_1$正交的矢量；这也就是，$\eta=\beta_2$是适合$|\eta|=1, (\eta,\alpha_1)=0$的所有矢量$\eta$中间，使$Q(\eta)$达到最大值$\lambda_2$的一个矢量．第三个基矢量$\gamma_3$是适合$|\zeta|=1$，并与$\alpha_1, \beta_2$正交的所有矢量$\zeta$中间，使$Q(\zeta)$达到最大值的一个矢量，等等．这个逐次最大问题可以通过具有三个不等轴$a>b>c>0$的椭球（以倒过来的形式，即极大换作极小）加以形象地描述．最短主轴c是最小直径；次主轴b是所有同最短主轴垂直的直径中间最短的一个，等等．

于是(36)式的系数$\lambda_1, \cdots, \lambda_n$可表征为某一极值问题的解，这个极值问题只依赖于$Q$，而不依赖于特殊的坐标系．在化简过程中，只有当第一个最大值（或者后面的某一最大值）是由两个或更多个长度为1的不同矢量ξ_0和η_0给出的时候才可能产生极值不确定的情况．即使在这种情况，我们仍可证明λ_i是唯一的（§10.4）．

这就证明了下面的主轴定理．

定理21 任意n个变量的实二次型，对于适当的标准正交基可以取为对角型(36)．

由定理1，这组新基底$\alpha_1^*, \cdots, \alpha_n^*$可以按照原基底$\varepsilon_1=(1, 0, \cdots, 0), \cdots, \varepsilon_n=(0, \cdots, 0, 1)$表示成$\alpha_i^*=\sum_j p_{ij}\varepsilon_j$．进一步，因为矢量$\alpha_1^*, \cdots, \alpha_n^*$是标准的而且是正交的，所以系数矩阵$P=(p_{ij})$是正交矩阵．象定理2那样，老坐标$x_1, \cdots, x_n$则可按照新坐标$x_1^*, \cdots, x_n^*$表示成$x_j=\sum_i x_i^* p_{ij}$；换句话说，我们已对二次型做了变量的正

交变换. 于是定理 21 的"坐标变换"形式的结果可改写成"点变换"的形式, 即

推论 1 n 个变量的任意齐次二次函数可以通过正交点变换化为对角型(36).

这两个结果都称为"主轴定理". 如果二次型用与它对应的矩阵来代替, 那么这个定理断言

推论 2 对任意实对称矩阵 A, 存在实正交矩阵 P 使得 $PAP^{\tau}=PAP^{-1}$ 是对角矩阵.

换句话说, 我们证明了任意实对称矩阵与对角矩阵相似. 同定理 4 相比较, 我们看出标准型(36)中的 $\lambda_1, \cdots, \lambda_n$ 恰好是矩阵 A 的全体特征值.

在平面上, 方程 $Q(\xi)=1$ 的标准型只不过是 $\lambda_1 x^2+\lambda_2 y^2=1$. 这类方程包括椭圆 $(\lambda_1 \geqslant \lambda_2>0)$ 或双曲线 $(\lambda_1>0>\lambda_2)$ 普通的标准方程, 这些系数确定了轴的长度. 在三维空间中, 对三个系数 $\lambda_1, \lambda_2, \lambda_3$ 有类似的解释. 如果这三个系数都是正的, 则轨迹 $Q=1$ 是一个椭球; 如果其中一个是负的, 则是单叶双曲面; 如果其中两个是负的, 则是双叶双曲面; 如果三个系数都是负的, 则根本没有轨迹. (注意符号差和秩在这里所起的作用.)

同定理 4 的推论相比较, 我们看出(对于对称矩阵 A)二次函数 XAX^{τ} 的主轴正是线性变换 $X \longmapsto XA$ 的特征矢量. 因此得到

推论 3 对于实对称矩阵 A, 线性变换 $X \longmapsto XA$ 有一组基底是由具有实特征值的正交特征矢量构成.

推论 4 每个非奇异实矩阵 A 可以表示成乘积 $A=SR$, 其中 S 是对称正定矩阵, R 是正交矩阵.

证明 我们已经知道(定理 20) AA^{τ} 是对称的, 并且是正定的. 根据主轴定理, 存在正交矩阵 P 使得 $P^{-1}AA^{\tau}P$ 是对角矩阵,

并且是正定的．于是对角线元素都是正的，通过求这些元素的平方根，我们得到一个正定对角矩阵 T，满足 $T^2=P^{-1}AA^{\tau}P$，因此正定对称矩阵 $S=PTP^{-1}$ 满足 $S^2=AA^{\tau}$．如果我们证明了 $R=S^{-1}A$ 是正交的，则有 $A=SR$，如推论所要求的，于是推论得证．而实际上，$RR^{\tau}=S^{-1}AA^{\tau}(S^{-1})^{\tau}=S^{-1}S^2(S^{-1})^{\tau}=S(S^{-1})^{\tau}=SS^{-1}=I$．这是因为对称矩阵 S 有性质 $(S^{-1})^{\tau}=S^{-1}$．

推论 5　设 A 是任意实对称矩阵，B 是任意正定(实)对称矩阵，那么存在一个实非奇异矩阵 P 使得 PAP^{-1} 和 PBP^{-1} 同时为对角矩阵．

我们把这个推论的证明留作习题．求满足 $A\xi_i=\lambda_i B\xi_i$ 的一组矢量 ξ_i，称为广义特征矢量问题，它的解在振动理论中起着基本作用．

习　　题

1．考虑实二次型 $ax^2+2bxy+cy^2$．

(a) 证明：在正交变换下 $a+c$ 和 b^2-ac 是不变量．

(b) 设 $\operatorname{ctg}2\alpha=\dfrac{a-c}{2b}$，证明：这个二次型可以通过正交变换

$$x=x'\cos\alpha-y'\sin\alpha,$$
$$y=x'\sin\alpha+y'\cos\alpha$$

化成对角型．

2．证明：每个实斜对称矩阵 A 具有形式 $A=P^{-1}BP$，其中 P 是正交矩阵，B^2 是对角矩阵．

3．按照已给出的方法，通过正交变换把下列二次型化为对角型：

(a) $5x^2-6xy+5y^2$，

(b) $2x^2+4\sqrt{3}\,xy-2y^2$．

4．把正交变换

$$3x_1=2y_1-y_2+2y_3,$$
$$3x_2=-y_1+2y_2+2y_3,$$
$$3x_3=2y_1+2y_2-y_3$$

作用到二次型 $9x_1^2-9x_2^2+18x_3^2$ 上，其结果 Q 是 y_1, y_2, y_3 的二次型，直接证明：矢量$\left(\frac{2}{3}, \frac{2}{3}, -\frac{1}{3}\right)$给出 Q 在 $y_1^2+y_2^2+y_3^2=1$ 上的最大值为 18．用微积分中的方法来检验．

5．在单位圆 $x=\cos\theta, y=\sin\theta$ 上考虑二次型 $ax^2+2bxy+cy^2$．证明它的极值是(参看习题 1)：

$$\frac{(a+c)\pm\sqrt{(a+c)^2-4\Delta}}{2}, \qquad \Delta=ac-b^2$$

6．证明：找不到含有有理元素的正交矩阵，把 xy 化为对角型．

7．保持 $x_1^2+x_2^2+x_3^2-x_4^2$ 不变的线性变换称为罗伦兹 (Lorentz) 变换．证明：矩阵P定义一个罗伦兹变换当且仅当 $P^{-1}=SP^{T}S=SP^{T}S^{-1}$，其中 S 是个特殊的对角矩阵，其对角线元素为 1, 1, 1, −1．

8．(a) 证明：如果 $A=SR$，其中 S 是对称矩阵、R 是正交矩阵，那么 $S^2=AA^{T}$．

*(b) 证明：只存在一个正定矩阵 S，它满足 $S^2=AA^{T}$．(提示：S^2 的任意特征矢量一定是 S 的一个特征矢量．)

*9．证明定理 21 的推论 5．(提示：把 XAX^{T} 看作具有内积 XBX^{T} 的欧几里得矢量空间的二次函数，再根据定理 20，把 B 写成 $B=PP^{T}$．)

§9.11 仿射群和欧几里得群之下的二次型

下面考虑含有坐标 $x_1, \cdots, x_n$ 的矢量 ξ 的任意非齐次二次函数

$$f(\xi)=\sum_i\sum_j x_i a_{ij} x_j+\sum_k b_k x_k+c \quad (i, j, k=1, \cdots, n). \tag{37}$$

这可以写成 $f(\xi)=XAX^{T}+BX^{T}+c$，其中 $A=(a_{ij})$是对称矩阵，$B=(b_1, \cdots, b_n)$是行矩阵．在单变量函数 $f=ax^2+bx+c$ 的简单情形中，我们看出，平移 $x=y+k$ 使二次项系数 a 保持不变，这是因为

$$\begin{aligned} f&=a(y+k)^2+b(y+k)+c \\ &=ay^2+(2ak+b)y+ak^2+bk+c. \end{aligned} \tag{38}$$

对 n 个变量的情形做类似的计算，平移 $X\longmapsto Y=X-K$（K 是行矩阵）给出

$$f(\xi)=(Y+K)A(Y+K)^{\tau}+B(Y+K)^{\tau}+c$$
$$=YAY^{\tau}+KAY^{\tau}+YAK^{\tau}+KAK^{\tau}+BY^{\tau}+BK^{\tau}+c.$$

因为乘积 YAK^{τ}（行矩阵×矩阵×列矩阵）是一个标量，所以它等于它的转置 $KA^{\tau}Y^{\tau}=KAY^{\tau}$，总起来有

$$f(\xi)=YAY^{\tau}+(2KA+B)Y^{\tau}+KAK^{\tau}+BK^{\tau}+c. \tag{39}$$

这确实与公式(38)类似. 这就证明了

引理 平移使二次函数 $f(\xi)$ 的齐次二次项部分对应的矩阵 A 保持不变.

另一方面，齐次线性变换 $X=YP$ 把 $f(\xi)$ 变到 $Y(PAP^{\tau})Y^{\tau}+(BP^{\tau})Y^{\tau}+c$. 在这个二次函数中，二次项的新矩阵是 PAP^{τ}，这恰好同单独是齐次二次型的变换情形一样.

现在通过具有方程 $X=YP+K$（P 是正交矩阵）的刚体运动来化简实函数 $f(\xi)$. 根据上述说明，只是 P 所对应的正交变换可以化简二次项的矩阵 A，确切地说，是齐次二次型对应的矩阵 A. 象在 §9.10 中那样，我们得到（用新的系数 b_i'）

$$f(\xi)=\lambda_1 z_1^2+\cdots+\lambda_n z_n^2+b_1' z_1+\cdots+b_n' z_n+c.$$

同非零的 λ_j 相联系的 b_j'，现在可以通过简单的“配方”法用平移 $y_j=z_j+\dfrac{b_j'}{2\lambda_j}$ 消去. 我们把变量置换一下，使得非零的 λ 项都移到前面，于是我们得到

$$f(\xi)=\lambda_1 y_1^2+\cdots+\lambda_r y_r^2+b_{r+1}' z_{r+1}+\cdots+b_n' z_n+c'.$$

如果这个函数的线性部分不正好是常数 c'，那么可以通过适当的平移和正交变换，象在定理 13 中那样，化为 dy_{r+1} 的形式. 这个变换并不影响前 r 个变量. 最后结果是下面形式中的一种

$$f(\xi)=\lambda_1 y_1^2+\cdots+\lambda_r y_r^2-dy_{r+1}, \tag{40}$$

$$f(\xi)=\lambda_1 y_1^2+\cdots+\lambda_r y_r^2+c'. \tag{41}$$

这里 $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_r$，所有 λ_i 都不为零，$d>0$.

定理 22 在所有刚体运动的欧几里得群之下，任意实二次型(37)等价于(40)或(41)式中的一种.

这些简化形式确实是实二次型在欧几里得群之下的标准型，但是证明比较困难，只概括叙述如下.

这些 λ_i 是(37)式相对应的矩阵 A 的全部特征值(见 § 9.10)；它们的唯一性(包括重数)将在 § 10.4 中证明. 特别是，(40)或(41)式中的平方项个数 r 是一个不变量；注意，r 也是 A 的秩，它在变换 $A \longrightarrow PAP'$ 之下是不变的. 利用微积分，可以对不变量 d 和 c' 作更为简单的直观描述. 我们考虑使得矢量

$$\operatorname{grad} f=\left(\frac{\partial f}{\partial y_1}, \cdots, \frac{\partial f}{\partial y_n}\right)$$

是零矢量的轨迹. 在(41)式的情况下，这个轨迹是子空间 $y_1=\cdots=y_r=0$，而 c' 是 $f(\xi)$ 在这个(不变)轨迹上的常数值. 在(40)式的情况下，这个轨迹是空的，因为 $\dfrac{\partial f}{\partial y_{r+1}}=d \neq 0$；而 d 可以表征为 $|\operatorname{grad} f|$ 的最小值；还可以证明，这个最小值在欧几里得群之下是不变量.

对于仿射变换 $X=YP+K$(其中 P 是非奇异的)，可以采用类似的处理方法. 在化简二次项部分为对角型时，象在 § 9.9 那样，所有系数都是 ± 1. 而线性部分可象 § 9.6 中那样去处理.

定理 23 n 个变量的任意实二次函数可以通过仿射变换(或者通过坐标仿射变换)化为下面形式中的一种.

$$y_1^2+\cdots+y_p^2-y_{p+1}^2-\cdots-y_r^2+c \quad (r \leqslant n), \tag{42}$$

$$y_1^2+\cdots+y_p^2-y_{p+1}^2-\cdots-y_r^2+y_{r+1} \quad (r<n). \tag{43}$$

因为平移不影响二次项，所以根据惯性律(定理 17)，秩 r 和

正项个数 p 一定是不变量.

从几何观点来看，每个二次函数 $f(\xi)=XAX'+BX'+c$ 定义一个图形或轨迹，它是由满足方程 $f(\xi)=0$ 的所有矢量 ξ 组成. 在二维空间中，根据这类二次方程画得的图形就是普通的二次曲线；在三维空间中，它是二次曲面；在一般情形中，它称为超二次曲面(或者二次超曲面). 一个仿射变换 $Y=XP+K$ 作用于这个曲面的方程也就相当于把同一个变换作用到图形的全部点上，并且我们称新的图形在给定的仿射变换之下等价于老的图形.

显然，上述的在等价之下二次函数分类的结论将给出对应图形的一个类似的分类. 然而，我们首先看出，方程 $f(\xi)=0$ 与同一方程的"数乘"积 $cf(\xi)=0$ 给出同样的轨迹. 这可以用来简化标准型，例如上面求得的 $y_1^2-y_2^2+c=0$. 当 $c\neq0$ 时，这个方程给出的轨迹同 $(c^{-1})y_1^2-(c^{-1})y_2^2+1=0$ 给出的轨迹一样；当 $c>0$ 时，通过仿射变换 $y_1=\sqrt{c}\,z_1$, $y_2=\sqrt{c}\,z_2$ 可以把它化为 $z_1^2-z_2^2+1=0$，而对于 $c<0$，变换 $y_i=\sqrt{-c}\,z_i$ 给出类似的结果 $z_2^2-z_1^2+1=0$. 一般地，用这个方法总可以把(43)式中出现的常数变为 1 或 0. 所以在实数域上的 n 维矢量空间中，任意超二次曲面在仿射群之下等价于下面方程中的一个所给出的轨迹：

$$y_1^2+\cdots+y_p^2-y_{p+1}^2-\cdots-y_r^2+1=0, \tag{44}$$

$$y_1^2+\cdots+y_p^2-y_{p+1}^2-\cdots-y_r^2+y_{r+1}=0, \tag{45}$$

$$y_1^2+\cdots+y_p^2-y_{p+1}^2-\cdots-y_r^2=0, \tag{46}$$

这里 $0\leqslant p\leqslant r\leqslant n$，在(45)的情形中 $r<n$.

在(44)式中，不同的二次型表示仿射不等价的轨迹，而在(45)式中，变换 $y_{r+1}\longmapsto -y_{r+1}$ 把 p 和 $r-p$ 互换，于是变换后的二次型与原来的等价.

例如，平面上满足 $r>0$ 的各种可能类型的轨迹是：

$$r=2$$

$x^2+y^2+1=0$	没有轨迹
$x^2-y^2+1=0$	双曲线
$-x^2-y^2+1=0$	圆
$\pm(x^2+y^2)=0$	一点
$x^2-y^2=0$	两条相交的直线

$$r=1$$

$\pm x^2+y=0$	抛物线
$x^2+1=0$	没有轨迹
$-x^2+1=0$	两条平行的直线
$x^2=0$	一条直线

特别要注意，不同的标准函数 x^2+y^2+1 和 x^2+1 给出相同的轨迹（即，这个图形一个点也没有）. 还有标准函数 x^2+y^2 和 $-x^2-y^2$, x^2+y 和 $-x^2-y$ 也是这样.

习　　题

1. 在欧几里得群之下，将下列二次型分类：
   (a) $4xz+4y^2+8y+8$.
   (b) $9x^2-4xy+6y^2+3z^2+2\sqrt{5}x+4\sqrt{5}y+12z+16$.

2. 在仿射群之下，将下列二次型分类：
   (a) $x^2+4y^2+9z^2+4xy+6xz+12yz+8x+16y+24z+15$.
   (b) $x^2-6xy+10y^2+2xz-20z^2-10yz-40z-17$.
   (c) $x^2+4z^2+4xz+4x+4z-6y+6$.
   (d) $-2x^2-3y^2-7z^2+2xy-8yz-6xz-4x-6y-14z-6$.

3. 证明：在二次函数 $XAX^\tau+BX^\tau+c$（其中 A 是非奇异矩阵）中的线性部分可以通过平移消去.

4. (a) 证明：非平凡实二次方程 $XAX^\tau=1$ 是一个旋转曲面当且仅当 A 具有二重特征值.

   (b) 描述二次方程 $xy+yz+zx=3$.

5. 把定理 23 中给出的二次函数的仿射分类推广到系数在任意满足 $1+1\neq 0$ 的域中的函数上.

6. (a) 列出三维空间中二次曲面的各种可能的仿射类型.

(b) 对每一类型给出简短的几何描述.

7. 在广义相似群(§ 9.4 末尾)之下,按(a)椭圆、(b)抛物线、(c)双曲线进行分类. 对每种情形求出全系不变量.

*8. 在刚体运动群之下,对 n 维欧几里得空间中的二次超曲面进行分类(利用定理 22).

9. 求在椭圆 $x^2+3y^2=3$ 中面积最大的内接六边形.

*§ 9.12 酉矩阵与埃尔米特矩阵

对于复数的情况,实二次型的正交变换是用某个埃尔米特型的酉变换来代替. 一个复数 $c=a+ib$ 定义为实数对(a, b),或定义为二维矢量空间 $\mathbf{R}^2$ 中具有分量(a, b)的矢量. 复数的模或绝对值$|c|$正好是实矢量的长度

$$|c|^2=|a+ib|^2=a^2+b^2=(a+ib)(a-ib)=cc^*, \quad (47)$$

这里 c^* 表示 c 的复共轭 $a-ib$. 同样道理,具有 n 个复分量$(c_1, \cdots, c_n)$(其中每个分量 $c_j=a_j+ib_j$)的矢量 γ 可以看作 $2n$ 维实空间中的具有 $2n$ 个分量$(a_1, b_1, \cdots, a_n, b_n)$的矢量. 这个实矢量的长度由下式的平方根给出

$$\begin{aligned}|(c_1, \cdots, c_n)|^2&=(a_1^2+b_1^2)+\cdots+(a_n^2+b_n^2)\\&=\sum_{j=1}^{n}(a_j+ib_j)(a_j-ib_j) \quad (48)\\&=c_1c_1^*+\cdots+c_nc_n^*.\end{aligned}$$

因为每个乘积 $c_jc_j^*=a_j^2+b_j^2\geqslant 0$,所以这个表达式具有严格的正性:这个实数和 $\sum_{j=1}^{n}c_jc_j^*$ 是正的,除非所有的 $c_j=0$. 表达式(48)很象通常的关于实矢量长度的毕达哥拉斯公式. 我们用(48)式作

为复行矢量 $K=(c_1,\cdots,c_n)$ 长度的定义. 公式 $\sum c_j c_j^*$ 可以写成矩阵形式 KK^*, 其中 K^* 是由矢量K的每个分量取共轭而得到的矢量.

定义 在复矢量空间 $\mathbf{C}^n$ 中,设 ξ 和 η 是具有坐标 $X=(x_1,\cdots,x_n)$ 和 $Y=(y_1,\cdots,y_n)$ 的矢量. 并引进内积

$$(\xi,\eta)=x_1y_1^*+\cdots+x_ny_n^*=XY^{*T}, \tag{49}$$

那么, ξ 的长度是 $|\xi|=(\xi,\xi)^{\frac{1}{2}}$.

同普通内积的情况几乎一样, 我们可以证明这个内积具有基本性质

线 性 $(c\xi+d\eta,\zeta)=c(\xi,\zeta)+d(\eta,\zeta)$.

斜对称性 $(\xi,\eta)=(\eta,\xi)^*$.

正 性 如果 $\xi\neq\mathbf{0}$, 则(ξ,ξ)是实数, 并且$(\xi,\xi)>0$.

由斜对称性显然可推出第二个因子的斜线性:

$$\begin{aligned}(\xi,c\eta+d\zeta)&=(c\eta+d\zeta,\xi)^*=c^*(\eta,\xi)^*+d^*(\zeta,\xi)^*\\&=c^*(\xi,\eta)+d^*(\xi,\zeta).\end{aligned}$$

所以

$$(\xi,c\eta+d\zeta)=c^*(\xi,\eta)+d^*(\xi,\zeta). \tag{50}$$

如果需要的话, 我们可以把线性、斜对称性和正性这些性质作为复数域上抽象矢量空间中内积(ξ,η)的公设. 那么这个空间被称为酉空间(比较 § 7.10 的欧几里得矢量空间).

两个矢量 ξ 和 η 如果满足$(\xi,\eta)=0$, 则称它们是正交的(记作 $\xi\perp\eta$). 根据斜对称性, $\xi\perp\eta$ 可推出 $\eta\perp\xi$. n维复矢量空间中一组(n 个)矢量 $\alpha_1,\cdots,\alpha_n$, 如果每个矢量的长度为 1, 且任意两个矢量是正交的:

$$|\alpha_1|=\cdots=|\alpha_n|=1, \quad (\alpha_i,\alpha_j)=0 \ (i\neq j) \tag{51}$$

则称这 n 个矢量是这个空间的标准酉基. 这样一组矢量一定是普

通意义下的基底．原来的基矢量 $e_1=(1,0,\cdots,0),\cdots,e_n=(0,\cdots,0,1)$ 构成一组标准酉基．根据 §7.11 的方法，我们还可以构造另外的标准酉基，并可证明

定理 24 *在酉空间中，任意 $m(<n)$ 个长度为 1 的相互正交的矢量构成这个空间标准酉基的一部分．*

特别是，如果 $\alpha_1,\cdots,\alpha_m$ 是非零正交矢量，$c_i=\dfrac{(\xi,\alpha_i)}{(\alpha_i,\alpha_i)}$，那么对任意 ξ，$\alpha_{m+1}=\xi-c_1\alpha_1-\cdots-c_m\alpha_m$ 都同矢量 $\alpha_1,\cdots,\alpha_m$ 正交．

一个 $n\times n$ 复矩阵 $U=(u_{ij})$，如果满足 $UU^{*\tau}=I$（其中 U^* 表示由 U 的每个元素取共轭而得到的矩阵），则称 U 为**酉矩阵**．这个条件显然等价于 $\sum_k u_{ik}u_{jk}^*=\delta_{ij}$，这里 δ_{ij} 是克罗内克尔符号（§9.4）．换句话说，U 的每一行矢量的长度为 1，U 的任意两个行矢量正交．这意味着，由 U 定义的 $\mathbf{C}^n$ 的线性变换把 $e_1,\cdots,e_n$ 变换到一组标准酉基．根据 §8.6 定理 9 的推论 6，它还等价于条件 $U^{*\tau}U=I$，这表明 U 的每一列矢量的长度为 1，U 的任意两个列矢量正交．

$\mathbf{C}^n$ 的任意线性变换 $X\longmapsto XA$ 把内积 $XY^{*\tau}$ 变换成 $XAA^{*\tau}Y^{*\tau}$．对所有的矢量 X 和 Y，这个新的内积等于原来的内积 $XY^{*\tau}=XIY^{*\tau}$ 当且仅当 $AA^{*\tau}=I$，即当且仅当 A 是酉矩阵．于是，一个矩阵 A 是酉矩阵当且仅当 A 所对应的线性变换 T_A 保持复内积 $XY^{*\tau}$ 不变．类似的论证指出，A 是酉矩阵当且仅当 T_A 保持长度 $(XX^{*\tau})^{\frac{1}{2}}$ 不变．几何上，如果酉空间的线性变换 T 保持长度不变 $|\xi T|=|\xi|$（因而也保持内积不变），那么称 T 为**酉变换**．n 维空间的所有酉变换组成的集合构成一个群，这个群与所有 $n\times n$ 酉矩阵组成的群同构．

下面，我们用"埃尔米特型"来代替二次型，它的最简单的例子是长度公式 $\sum x_ix_i^*$．一般地，**埃尔米特型**是一个含有复系数

h_{ij}的表达式

$$\sum_{i,j=1}^{n} x_i h_{ij} x_j^* = XHX^{*\tau}, \quad H = (h_{ij}), \tag{52}$$

其中系数矩阵H具有性质 $H^{*\tau}=H$. 这种类型的矩阵H称为埃尔米特矩阵. 当元素 h_{ij} 都是实数时, 埃尔米特矩阵就是对称矩阵. 埃尔米特型(52)可以看作对于某基底的坐标为 $x_1, \cdots, x_n$的矢量ξ的函数 $h(\xi)=XHX^{*\tau}$. 这个函数的值 $XHX^{*\tau}$ 总是实数. 为证明这一结论,只须证明这个数值等于它的共轭(或者证明等于它的共轭转置). 而因为H是埃尔米特矩阵,所以有

$$(XHX^{*\tau})^{*\tau}=(X^*H^*X^{**\tau})^{\tau}=X^{\tau\tau}H^{*\tau}X^{*\tau}=XHX^{*\tau};$$

满足断言.

酉变换 $Y=XU$, $X=YU^{-1}=YU^{*\tau}$ 应用到埃尔米特型,得到

$$XHX^{*\tau}=(YU^{-1})H(YU^{*\tau})^{*\tau}=YU^{-1}H(UY^{*\tau})$$
$$=Y(U^{-1}HU)Y^{*\tau}.$$

这个系数矩阵 $U^{-1}HU$ 还是埃尔米特矩阵, 这因为由于 $U^{-1}=U^{*\tau}$, 有

$$(U^{-1}HU)^{*\tau}=U^{*\tau}H^{*\tau}(U^{-1})^{*\tau}=U^{-1}HU.$$

如果我们用坐标变换, 变换到一个新的标准酉坐标系, 那么可以得到同样的效果, 因为这样的变换通过方程 $Y=XU$ (其中U是酉矩阵)给出ξ的新坐标Y与老坐标的关系.

用这种变换的说法, 我们可以把任意埃尔米特型变换到主轴上. 这个新主轴是根据逐次最大的性质而选取的, 这恰好同二次型在正交变换下化到主轴的讨论一样. 第一轴α_1取作长度为1的矢量,并使$h(\xi)$在$|\xi|=1$上取最大值;然后我们根据定理24可以求出包含α_1在内的标准酉基. 对于这组基底, 交叉乘积项 $x_1x_j^*$ $(j\neq 1)$又被去掉. 因为埃尔米特型的值总是实的, 所以逐次最大值λ_i都是实数. 这一过程证明了下面的主轴定理.

定理 25 任意埃尔米特型 $XHX^{*\prime}$ 可以通过酉变换 $Y=XU$ 化成实对角型

$$YHY^{*\prime}=\lambda_1 y_1 y_1^*+\lambda_2 y_2 y_2^*+\cdots+\lambda_n y_n y_n^*. \qquad (53)$$

这个定理可以平移到埃尔米特型的矩阵 H 上，即

定理 26 对于每个埃尔米特矩阵 H，存在一个酉矩阵 U，使得 $U^{-1}HU=U^{*\prime}HU$ 是实对角矩阵。

用第十章的方法还可证明(53)式的系数 λ_i 是唯一的。

习 题

1. 下列矩阵中哪些是酉矩阵？哪些是埃尔米特矩阵？

$$\begin{pmatrix}\frac{1+i}{2} & \frac{1-i}{2}\\ \frac{1-i}{2} & \frac{1+i}{2}\end{pmatrix},\quad \begin{pmatrix}3 & 1-i\\ 1+i & \sqrt{2}\end{pmatrix},\quad \begin{pmatrix}1 & i\\ i & 1\end{pmatrix}.$$

2. 求与 $\left(\frac{1}{2},\ \frac{i}{2},\ \frac{1+i}{2}\right)$ 正交的矢量组成的子空间的一组标准酉基。

3. 证明：(h_{ij}) 是埃尔米特矩阵当且仅当对所有的 i, j 有 $h_{ij}^*=h_{ji}$。

4. 证明：如果 ω 是 n 次本原单位根，那么 $n^{-\frac{1}{2}}(\omega^{ij})\ (i, j=1, \cdots, n)$ 是酉矩阵。

5. 证明：对任意实数 θ，复矩阵

$$\begin{pmatrix}\mathrm{ch}\theta & i\,\mathrm{sh}\theta\\ -i\,\mathrm{sh}\theta & \mathrm{ch}\theta\end{pmatrix}$$

是酉矩阵，计算它的特征值和特征矢量。

6. 证明：所有 $n\times n$ 酉矩阵构成一个群（酉群），这个群与所有 $2n\times 2n$ 实正交矩阵组成的群的一个子群同构。

7. 证明：埃尔米特内积 (ξ, η) 满足线性、斜对称性和正性。

8. 详细证明关于标准酉基的定理 24。

9. 证明：一个单项矩阵是酉矩阵当且仅当它的所有非零元素的绝对值是 1。

10. 对于两个变量的埃尔米特型，它在 $x=0, y=1$ 点处取最大值，证明类似于§9.10引理的结果。（提示：把每个变量分成实部和虚部。）

*11. 给出埃尔米特型主轴定理的详细证明.

12. 通过关于 x 和 y 的酉变换，把埃尔米特型 xy^*+x^*y 化成对角型.（提示：考虑相应的实二次型.）

13. 在酉群之下，把 $zz^*-2ww^*+2i(zw^*-wz^*)$ 化成对角型.

14. 证明：任意实斜对称矩阵 A 具有一组由复特征矢量组成的基底，这些特征矢量对应的特征值都是纯虚数.（提示：证明 iA 是埃尔米特矩阵.）

15. 证明：任意酉矩阵的谱位于复平面的单位圆上.

16. 证明：复矩阵 C 是正定的埃尔米特矩阵当且仅当 $C=PP^{*T}$，其中 P 是某非奇异矩阵.

17. 证明：埃尔米特矩阵是正定的当且仅当它的所有特征值都是正的.

*§9.13 仿射几何

仿射几何是研究在仿射群之下图形不变的性质．正如欧几里得几何是研究在欧几里得群之下图形不变的性质一样．作用在有限维矢量空间 V 上的仿射群，象(11)式那样，是由 V 的所有把 V 的点（或矢量）ξ 变到点

$$\xi H=\eta=\xi T+\kappa \tag{54}$$

的变换 H 组成的，这里 κ 是一固定矢量，T 是 V 的一固定非奇异线性变换．我们假定 V 是域 F 上的矢量空间，其中域 F 满足条件 $1+1\neq 0$（例如 F 不是域 $\mathbf{Z}_2$）.

在仿射几何中，正如在欧几里得几何中一样，任意两个点 α 和 β 是等价的，这因为平移 $\xi\longmapsto\xi+(\beta-\alpha)$ 把 α 变到 β．这就把仿射几何与 V 的矢量几何（在全线性群之下）区别开来，在仿射几何中原点 O 起着 V 中零矢量 $\mathbf{0}$ 的特殊作用．当考虑在仿射群之下保持不变的性质时，我们通常把矢量空间称作*仿射空间*.

在平面解析几何中，联结两点 (x_1, y_1) 和 (x_2, y_2) 的直线的方程是

$$y-y_1=\frac{y_2-y_1}{x_2-x_1}(x-x_1),\quad x_2\neq x_1.$$

引入参数 $t=\frac{x-x_1}{x_2-x_1}$，则我们得到 $y=y_1+t(y_2-y_1)$，$x=x_1+t(x_2-x_1)$；换句话说，这直线具有参数方程

$$x=(1-t)x_1+tx_2,\quad y=(1-t)y_1+ty_2, \tag{55}$$

它可以写成矢量形式 $(x,y)=\xi=(1-t)\xi_1+t\xi_2$. 几何上，(55)的点 (x,y) 是把 (x_1,y_1) 和 (x_2,y_2) 两点之间的线段分成比例为 $t:(1-t)$ 的点. 当 $t=\frac{1}{2}$，这个点就是中点.

在任意仿射空间中，把 α 和 β 之间的"线段"分成比例为 $t:(1-t)$ 的点定义为

$$\gamma=(1-t)\alpha+t\beta. \tag{56}$$

当 $\alpha\neq\beta$，联结 α 和 β 的(仿射)直线 $\overline{\alpha\beta}$ 定义为满足(56)式($t\in F$)的所有点 γ 组成的集合.

定理 27 任意非奇异仿射变换把直线变到直线.

证明 把(56)代入(54)中，我们有

$$\begin{aligned}\gamma H=\gamma T+\kappa&=(1-t)\alpha T+t\beta T+\kappa\\&=(1-t)(\alpha T+\kappa)+t(\beta T+\kappa)\\&=(1-t)(\alpha H)+t(\beta H).\end{aligned}$$

因此 H 把通过 α 和 β 的仿射直线 $\overline{\alpha\beta}$ 变到通过 αH 和 βH 的仿射直线. 证毕

如果 $\gamma=(1-t)\alpha+t\beta$ 和 $\delta=(1-u)\alpha+u\beta$ 是 $\overline{\alpha\beta}$ 上任意两个不同的点，那么因为

$$(1-v)\gamma+v\delta=(1-t+vt-vu)\alpha+(t-vt+vu)\beta$$

所以 $\overline{\alpha\beta}$ 包含 $\overline{\gamma\delta}$ 的每个点. 反过来也可以类似地证明 $\overline{\gamma\delta}$ 包含 $\overline{\alpha\beta}$ 的每个点，因此 $\overline{\alpha\beta}=\overline{\gamma\delta}$. 也就是说，直线可用它上面的任意两点确定.

一个普通平面有时可用平直性质来表征：平面如果包含任意

两点，则它必包含通过这两点的整个直线．我们可以用这种性质把V的仿射子空间定义为V中具有下面性质的任意子集合M：如果α和β在M中，则整个直线$\overline{\alpha\beta}$也在M中．显然，仿射变换把仿射子空间映上到仿射子空间．更进一步，V的仿射子空间恰恰是按下述意义平移V的矢量子空间而得到的一个子空间．

定理 28 *如果M是V的任意仿射子空间，那么存在V的一个线性子空间S和矢量κ，使得M是由所有点$\xi+\kappa$组成，其中$\xi\in S$．反过来，任意S和κ可以按这种方法确定一个仿射子空间$M=S+\kappa$．*

证明 设κ是M中任意一点，并定义S是所有矢量$\alpha-\kappa$组成的集合（其中$\alpha\in M$）；换句话说，S是按$-\kappa$平移M而得到的．显然，M具有所要求的按S和κ表达的形式．剩下只须证明S是一个矢量子空间．因为直线平移到直线，所以对于M的假设保证S也有同样的性质：联结S的任意两个矢量的直线仍在S中．对于S中任意矢量α，联结$O(\in S)$和α的直线在S中，因此S包含所有"数乘"积$c\alpha$．如果S包含α和β，那么它必包含2α和2β以及联结它们的整个直线$\xi=2\alpha+t(2\beta-2\alpha)$．（画一个图！）特别当$t=\frac{1}{2}$时，$S$包含$\xi=2\alpha+(\beta-\alpha)=\beta+\alpha$，也就是包含给定矢量的和．于是，我们就证明了$S$在加法和数乘运算之下是封闭的，因此它是矢量子空间，满足要求． 证毕

$F=\mathbf{Z}_2$的情形是个例外：三个矢量组$(0,0)$，$(1,0)$和$(0,1)$是一个"平面"，当它包含任意两点α和β，就必包含所有$(1-t)\alpha+t\beta$；但是这三矢量组并不是一个仿射子空间．

逆定理也不难证明．它换种说法就是，仿射子空间正好是在矢量加法群之下矢量子空间的陪集．特别是，仿射直线是一维矢量子空间的陪集（在平移之下）．

上述结果包含着仿射几何中另一个概念：平行性.

定义 仿射空间V的两个子集合S和S^*称为是平行的当且仅当存在V的一个平移$L:\xi\longmapsto\xi+\lambda$把$S$映上到$S^*$.

定理 29 V的任意仿射变换把平行集合映射到平行集合.

证明 设S和$S^*=S+\lambda$是给定的平行集合，设U和U^*分别是它们在$H:\xi\longmapsto\xi T+\kappa$之下的变换式. 定理断言，$U^*$是由所有$\xi+\mu$组成的集合，其中$\xi\in U$，$\mu$是某一固定的平移矢量. 根据定义，$U^*$是由所有$(\sigma+\lambda)T+\kappa=(\sigma T+\kappa)+\lambda T$组成的集合，其中$\sigma\in S$. 并且$U$是由所有$\xi=\sigma T+\kappa$组成的集合，其中$\sigma\in S$. 置$\mu=\lambda T$，显然定理的结论成立. 证毕

关于在仿射群(实数域$\mathbf{R}$上)之下的等价性在初等几何上有很多有趣的应用. 在仿射群之下，任意两个三角形是等价的. 为了证明这一点，只须证明任意三角形$\alpha\beta\gamma$等价于以$O=(0,0)$，$\beta_0=(2,0)$，$\gamma_0=(1,\sqrt{3})$为顶点的特殊的等边三角形(见图2). 通过平移，顶点α可以移动到原点O，其他两顶点移到位置β'和γ'. 因为矢量β'和γ'是线性无关的，则存在一个线性变换$x\beta'+y\gamma'\longmapsto x\beta_0+y\gamma_0$，它把$\beta'$变到$\beta_0$，把$\gamma'$变到$\gamma_0$. 这个变换与平移的乘积将把$\alpha\beta\gamma$变到$O\beta_0\gamma_0$，正如所求. 因此这两个三角形是等价的.

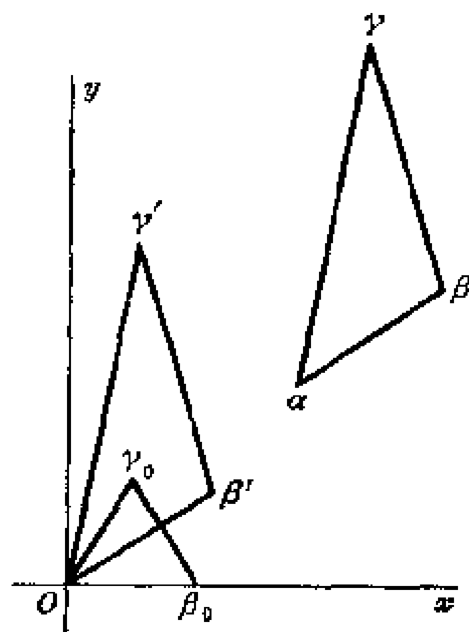


图 2

于是每个三角形等价于一个等边三角形. 而等边三角形的三条中线根据对称性它们一定交于一点(即重心). 但是，仿射变换把中点变到中点，因此它把中线变到中线. 这就证明了一个初等

定理：任意三角形三中线交于一点．再有，我们可以很容易地证明，等边三角形中线的交点按比例 1:2 把中线分成两部分，因此对任意三角形同样的结论成立．

此外，任意椭圆与一个圆仿射等价．但是一个圆的通过圆心的任意直径在两个端点上有互相平行的切线，而且平行于这两条切线的共轭直径把平行于给定直径的所有弦平分．可以推出，对任意椭圆，这两个性质同样成立，这是因为仿射变换把平行线变为平行线，把切线变为切线(但应注意，椭圆中的共轭直径不一定互相垂直，图 3)．

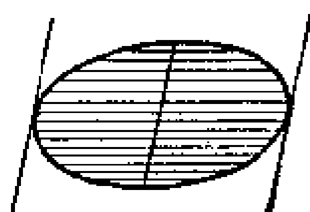


图 3

附录 形心与重心坐标 按给定的比例分割一线段的分点(56)是形心概念的特殊情形．在V中给出 $m+1$ 个点 $\alpha_0, \cdots, \alpha_m$，在$F$中给出 $m+1$ 个元素 $x_0, \cdots, x_m$，使得 $x_0+\cdots+x_m=1$，具有权 $x_0, \cdots, x_m$ 的点 $\alpha_0, \cdots, \alpha_m$ 的形心定义为点

$$\xi=x_0\alpha_0+\cdots+x_m\alpha_m, \quad x_0+\cdots+x_m=1. \tag{57}$$

(更一般地，对任意 $w=w_0+\cdots+w_m\neq 0$，具有权 $w_0, \cdots, w_m$ 的点 $\alpha_0, \cdots, \alpha_m$ 的形心仍用(57)来定义，这时 $x_i=\frac{w_i}{w}$.)

如果H是形为(54)的任意仿射变换，那么

$$\begin{aligned}\xi H &= (x_0\alpha_0+\cdots+x_m\alpha_m)T+\kappa\\ &= x_0(\alpha_0T)+\cdots+x_m(\alpha_mT)+(\sum x_i)\kappa\\ &= x_0(\alpha_0H)+\cdots+x_m(\alpha_mH).\end{aligned}$$

换句话说，仿射变换把形心变换到形心，其权不变．

定理 30 仿射子空间M包含它的点的所有形心．

我们通过对(57)式的点数 $m+1$ 用归纳法来证明这个定理．当 $m=0$，可直接得出结论．当 $m=1$，α_0 和 α_1 的形心恰好是通过

α_0 和 α_1 的直线上的一个点，因此根据定义，这个形心在M中．假定 $m>1$，并把 ξ 看作形为(57)的点．那么有某个系数，比如说是 x_m 不等于 1．设 $t=x_0+\cdots+x_{m-1}$，则 $x_m=1-t$，$t\neq 0$，点 $\beta=\frac{x_0}{t}\alpha_0+\cdots+\frac{x_{m-1}}{t}\alpha_{m-1}$ 是 $\alpha_0,\cdots,\alpha_{m-1}$ 的形心，由归纳法假设，它在M中．进一步，$\xi=t\beta+(1-t)\alpha_m$ 是在联结 $\beta\in M$ 和 $\alpha_m\in M$ 的直线上，因此 ξ 在 M 中，如断言所述．

形心可用来描述由给定的点集 $\alpha_0,\cdots,\alpha_m$ 张成的子空间 M，如下所述．

定理 31 V 中 $m+1$ 个点 $\alpha_0,\cdots,\alpha_m$ 的所有形心(57)组成的集合是一个仿射子空间 M．这个子空间M 包含每个点 α_i，并且它包含在任意仿射子空间N中，而N包含所有的点 $\alpha_0,\cdots,\alpha_m$.

证明 设(57)表示的 ξ 和

$$\eta=y_0\alpha_0+\cdots+y_m\alpha_m,\quad y_0+\cdots+y_m=1 \tag{57'}$$

是任意两个形心，则

$$\begin{aligned}&(1-t)\xi+t\eta\\&=[(1-t)x_0+ty_0]\alpha_0+\cdots+[(1-t)x_m+ty_m]\alpha_m\end{aligned}$$

也是 $\alpha_0,\cdots,\alpha_m$ 的形心，这是因为全体系数 $(1-t)x_i+ty_i$ 的和是 1．因此M的确是一个仿射子空间．显然它包含每个 α_i．另一方面，根据定理 30，包含所有 α_i 的任意仿射子空间 N，一定包含整个M． 证毕

如果m个矢量 $\alpha_1-\alpha_0,\cdots,\alpha_m-\alpha_0$ 是线性无关的，那么称 $m+1$ 个点 $\alpha_0,\cdots,\alpha_m$ 是仿射无关的．对于一个仿射变换 H，我们有 $(\alpha_i-\alpha_0)T=\alpha_iH-\alpha_0H$，因此非奇异仿射变换把仿射无关的点变到仿射无关的点．在这个仿射无关的定义中，起点 α_0 起着特殊的作用．下面的结果表明，仿射无关性不依赖于起点的选取．

定理 32 $m+1$ 个点仿射无关当且仅当由 $\alpha_0, \cdots, \alpha_m$ 张成的仿射子空间 M 中的每一点 ξ，作为 $\alpha_0, \cdots, \alpha_m$ 的形心(57)，有唯一的表达式.

证明 假设点 $\alpha_0, \cdots, \alpha_m$ 仿射无关，而 M 中某一点 ξ，作为形心有两种表达式 $\xi=\sum x_i\alpha_i$，$\xi=\sum x_i'\alpha_i$，并满足 $\sum x_i=1=\sum x_i'$. 那么

$$x_0'-x_0=(x_1-x_1')+\cdots+(x_m-x_m'),$$

而且零矢量 $\mathbf{0}=O$ 具有表达式

$$\mathbf{0}=\sum_{i=0}^{m}(x_i-x_i')\alpha_i=\sum_{i=1}^{m}(x_i-x_i')\alpha_i-(x_0'-x_0)\alpha_0$$

$$=\sum_{i=1}^{m}(x_i-x_i')(\alpha_i-\alpha_0)$$

因为矢量 $\alpha_1-\alpha_0, \cdots, \alpha_m-\alpha_0$ 是线性无关的，我们推出结论：对于 $i=1, \cdots, m$ 有 $x_i=x_i'$. 因为 $x_0=1-(x_1+\cdots+x_m)$，所以我们还有 $x_0=x_0'$. 于是，ξ 作为形心的表达式是唯一的.

其次，假设点 $\alpha_0, \cdots, \alpha_m$ 仿射相关. 那么存在一个线性关系 $\sum c_i(\alpha_i-\alpha_0)=\mathbf{0}$，其中系数 c_i 不全为零，比如说 $c_1\neq 0$. 通过做除法，我们可以假定 $c_1=1$. 则有

$$\alpha_1=-c_2\alpha_2-\cdots-c_m\alpha_m+(c_2+\cdots+c_m+1)\alpha_0,$$

在 α_1 的表达式中，这些系数之和是 1，但是 α_1 还有另一种表示 $\alpha_1=1\cdot\alpha_1$，因此 α_1 作为形心的表示不是唯一的. 证毕

当点 $\alpha_0, \cdots, \alpha_m$ 是仿射无关时，由 $\alpha_0, \cdots, \alpha_m$ 张成的空间中的每一点可表成(57)，在表达式(57)中出现的标量 $x_0, \cdots, x_m$ 称为 ξ 相对于 $\alpha_0, \cdots, \alpha_m$ 的重心坐标. 注意，这些坐标中任意 m 个可以确定剩下的一个坐标，这是由于 $x_0+\cdots+x_m=1$.

习 题

1. 对下列每对点，求两点联线的参数方程，并把这直线表示成形式 $S_1+\lambda$（也就是求空间 S_1）.

(a) (2, 1) 和 (5, 0),

(b) (1, 3, 2) 和 (−1, 7, 5),

(c) (1, 2, 3, 4) 和 (4, 3, 2, 1).

2. 把通过 (1, 3) 和 (4, 2) 两点的直线表示成形式 $S+\lambda$，其中 λ 有四种不同的选择. 画出图形.

3. 证明：通过不在一直线上的三个矢量 α, β, γ，存在一个且只存在一个二维仿射子空间（一个平面！）. 证明：这个平面上的矢量可表示成形式 $\xi=\alpha+s(\beta-\alpha)+t(\gamma-\alpha)$，这里 s 和 t 是变量.

4. 求通过下列每三点的平面（如果存在的话）的参数方程（按照习题 3 的形式）：

(a) (1, 3, 2), (4, 1, −1), (2, 0, 0),

(b) (1, 1, 0), (1, 0, 1), (0, 1, 1),

(c) (2, −1, 3), (1, 1, 1), (3, 0, 4).

5. 在习题 4 的每个小题中，求通过原点的平行平面的基底.

*6. 证明：定理 28 在除去 $\mathbf{Z}_2$ 的每个域上都成立.

7. 只用有关的定义证明：任意仿射变换把中点变到中点.

8. 证明：每个平行四边形与正方形仿射等价.

9. 用仿射方法证明：平行四边形的两条对角线总是彼此平分.

10. (a) 求 $\mathbf{R}^2$ 的仿射变换，它把以 (0, 0), (0, 1), (1, 0) 为顶点的三角形变换到以 (1, 0), (−1, 0), $(0, \sqrt{3})$ 为顶点的等边三角形.

(b) 如果第一个三角形是以 (1, 1), (1, 2), (3, 3) 为顶点，那么所求的仿射变换是什么？

11. 用仿射方法证明：在梯形中，两条对角线和两平行边中点联线通过一点.

12. 证明：任意平行六面体与正立方体仿射等价.

13. 证明：任意平行六面体的四条对角线有公共的中点（它是重心）.

14. (a) 证明：在任意域 F 上，任意两个三角形在仿射群之下是等价的.

*(b) 证明：如果域 F 中 $1+1\neq 0, 1+1+1\neq 0$，那么任意三角形的中线交于一点。

15. 证明：矢量空间 V 的一一变换 T 是仿射变换当且仅当 $\gamma=(1-t)\alpha+t\beta$ 总可推出 $\gamma T=(1-t)\alpha T+t(\beta T)$。

16. 证明：如果仿射子空间 M 是由 $m+1$ 个仿射无关的点 $\alpha_0, \cdots, \alpha_m$ 张成，那么 M 同一个 m 维矢量子空间平行。

17. 根据定义，F^n 中的超平面是 $n-1$ 维仿射子空间。

(a) 证明：其坐标满足线性方程 $a_1x_1+\cdots+a_nx_n=c$ 的所有矢量 ξ 组成的集合是一个超平面，这里假定 $a, \cdots, a_n$ 不全为零。

(b) 反过来证明：每个超平面都满足这样的方程。

(c) 求通过四个点 $(1,0,1,0)$、$(0,1,0,1)$、$(0,1,1,0)$、$(1,0,0,1)$ 的超平面方程。

18. 设 $\alpha_0, \cdots, \alpha_n$ 是 n 维矢量空间 V 的 $n+1$ 个仿射无关的点，且设 $\beta_0, \cdots, \beta_n$ 是 V 中任意 $n+1$ 个点。证明：存在一个且只存在一个 V 的仿射变换把每个 α_i 变到 β_i。

19. 证明：如果仿射子空间 M 是由 $m+1$ 个仿射无关的点 $\alpha_0, \cdots, \alpha_m$ 张成，并且由 $r+1$ 个仿射无关的点 $\beta_0, \cdots, \beta_r$ 张成，那么 $m=r$。

*§9.14 射影几何

在实仿射平面中，任意两个点在唯一的一条直线上，任意两条不平行的直线交于唯一的一点。我们现在来构造实射影平面，在这个平面上，

（i）任意两个不同的点在唯一的一条直线上。

（ii）任意两条不同的直线交于唯一的一点。相关联的性质(i)和(ii)显然在下述意义下彼此对偶，即互换“点”和“直线”二词，并在术语上稍加变化，就可把性质(i)变为性质(ii)，也可把性质(ii)变为性质(i)。

构造实射影平面 $P_2=P_2(\mathbf{R})$ 的一种方法如下所述。取实数域 $\mathbf{R}$ 上一个三维矢量空间 V_3，并把 V_3 的一维矢量子空间（不是仿射

子空间）S 称为 P_2 的一个"点"，把 V_3 的二维子空间 L 称为 P_2 的"直线"．进一步，我们说点 S 在直线 L 上当且仅当子空间 S 包含在子空间 L 中．

我们来证明 $P_2(\mathrm{R})$ 的"点"和"直线"满足(i)和(ii)．如果 S_1 和 S_2 分别是由矢量 α_1 和 α_2 张成的一维子空间，那么 $S_1\neq S_2$ 当且仅当 α_1 和 α_2 是线性无关．于是 S_1 和 S_2 所在的唯一直线 L 就是由 α_1 和 α_2 张成的二维矢量子空间，这就证明了性质(i)．其次，如果直线（二维子空间）L_1 和 L_2 是不同的，那么子空间 L_1+L_2（是 L_1 与 L_2 的线性和）必具有较高的维数，于是它是整个三维空间 V_3．因此根据§7.8 定理 17，有

$$\begin{aligned}\dim(L_1\cap L_2)&=\dim L_1+\dim L_2-\dim(L_1+L_2)\\&=2+2-3=1,\end{aligned}$$

所以一维子空间 $L_1\cap L_2$ 是同时位于 L_1 和 L_2 的唯一的一点．这就证明了性质(ii)．

为了在射影平面 $P_2=P_2(\mathrm{R})$ 中得到适当的射影坐标，取 V_3 是由所有 3-实数组 (x_1, x_2, x_3) 组成的空间 R^3．那么每个非零的 3-数组 (x_1, x_2, x_3) 确定 P_2 的一点 S；当 $c\neq 0$ 时，3-数组 (x_1, x_2, x_3) 和 (cx_1, cx_2, cx_3) 确定同一个点 S．我们把这些 3-数组等同起来

$$(x_1, x_2, x_3)=(cx_1, cx_2, cx_3),\quad c\neq 0$$

并称这些 3-数组为点 S 的齐次坐标．因为 V_3 的任意二维子空间 L 可以描述为一个齐次线性方程的解矢量组成的集合，所以 P_2 的直线 L 是那些齐次坐标满足方程

$$a_1x_1+a_2x_2+a_3x_3=0,\quad (a_1, a_2, a_3)\neq(0, 0, 0) \tag{58}$$

的点的轨迹．我们可称 (a_1, a_2, a_3) 为直线 L 的齐次坐标．显然，当 $c\neq 0$ 时，坐标 (a_1, a_2, a_3) 和 (ca_1, ca_2, ca_3) 确定同一条直线．

实射影平面有非常简单的几何表示．点 S 的齐次坐标 $(x_1, x_2,$

x_3)，通过乘上$(x_1^2+x_2^2+x_3^2)^{-\frac{1}{2}}$可以标准化，于是新坐标$(y_1, y_2, y_3)$满足$y_1^2+y_2^2+y_3^2=1$并在单位球面上，在这个球面上的两个对径点$(y_1, y_2, y_3)$和$(-y_1, -y_2, -y_3)$确定$P_2$的同一点．换句话说，$P_2$的点可以通过把直径的两端点（在球面上）看作一个点而得到．因为V_3的任意二维矢量子空间L沿一个大圆截割单位球，所以我们可以说，P_2的一条直线是由单位球的一个大圆上的全体对径点对组成．于是这又表明，两条射影直线（两个大圆）交于一个射影点（球面上一对对径点）．

可以用同样的方法在任意域F上定义一个射影平面$P_2(F)$．在任何情形下，显然有每个一维矢量子空间(cx_1, cx_2, cx_3)（其中$x_3\neq0$）与仿射平面$x_3=1$恰恰交于一点$\left(\frac{x_1}{x_3}, \frac{x_2}{x_3}, 1\right)$；这个比$\left(\frac{x_1}{x_3}, \frac{x_2}{x_3}\right)$称为射影点$(cx_1, cx_2, cx_3)$的非齐次坐标．但是$x_3=0$的轨迹是一条射影直线，它称为“在无穷远处的直线”．可以验证，射影平面P_2的每条直线

$$L: a_1x_1+a_2x_2+a_3x_3=0$$

或者是“在无穷远处的直线”（当$a_1=a_2=0$），或者是仿射平面的一条直线$a_1\left(\frac{x_1}{x_3}\right)+a_2\left(\frac{x_2}{x_3}\right)+a_3=0$，加上“在无穷远处的直线”上的一个点$(a_2, -a_1, 0)$．

在任意域F上可以构造n维射影空间P．本质的一步是从高一维的矢量空间$V=F^{n+1}$开始．那么$P=P_n(F)$可以描述如下：P的点是V的一维子空间S，P的m维子空间是由位于V的某个$m+1$维矢量子空间L的所有点S（P的点）组成的集合．显然，每个这样的子空间本身与由这个$m+1$维矢量空间L按同样方法确定的m维射影空间P_m同构．如果V表示为F的$n+1$-元素组构成的空间（用对于给定基底的坐标表示），那么P_n的每个点S可以

用 $n+1$ 个齐次坐标$(x_1, \cdots, x_{n+1})$给出，并且坐标$(cx_1, \cdots, cx_{n+1})$（其中 $c\neq 0$）与坐标$(x_1, \cdots, x_{n+1})$确定同一个点.

$P=P_n(F)$中的超平面（$n-1$ 维子空间）仍是由一个齐次方程

$$a_1x_1+\cdots+a_{n+1}x_n=0, \quad (a_1, \cdots, a_{n+1})\neq(0, \cdots, 0) \qquad (59)$$

给出的轨迹. 数组$(a_1, \cdots, a_{n+1})$可以看作超平面的齐次坐标；射影空间P和它的对偶射影空间（这个空间的点是P的超平面）之间的关系，同矢量空间V和它的对偶空间V^*之间的关系一样. 根据§7.7定理13（关于齐次线性方程组解集合的维数的定理），我们得到，r 个象(59)那样的线性无关的方程组可以确定一个 $n-r$ 维射影子空间.

设$T:V\to V$ 是一个非奇异线性变换. 我们知道（§8.6定理10的推论2），T 把V 的每个一维子空间 S 变换到V 的一个一维子空间 S^*. 因此T诱导出射影空间P的点的变换 $S\longmapsto S^*=ST^*$，这个变换 T^* 把射影子空间变换到射影子空间，其维数保持不变. 我们称 T^* 是P的射影变换. 如果T_1和T_2是V的两个这样的线性变换，那么乘积T_1T_2诱导出P上的一个变换$(T_1T_2)^*$，它是这两个诱导出的变换的乘积$T_1^*T_2^*$. 因此所有射影变换组成的集合构成一个群，即 n 维射影群，并且对应 $T\longmapsto T^*$ 是 $n+1$ 维全线性群到域F上 n 维射影群上的一个同态.

相对于V中给定的坐标系，线性变换T是由一个非奇异$(n+1)\times(n+1)$矩阵(a_{ij})确定的. 那么变换 T^* 把具有齐次坐标$(x_1, \cdots, x_{n+1})$的点变换到具有齐次坐标$(y_1, \cdots, y_{n+1})$的点，其中

$$y_j=x_1a_{1j}+\cdots+x_{n+1}a_{n+1,j} \quad (j=1, \cdots, n-1). \qquad (60)$$

定理33 $(n+1)\times(n+1)$矩阵 A 确定 P_n 的恒等射影变换T^*当且仅当 A 是单位矩阵 I 的数乘积 cI（其中 $c\neq 0$）.

证明 在公式(60)中，如果 $A=cI$，则 $y_j=cx_j$；即齐次坐标$(x_1, \cdots, x_{n+1})$和$(cx_1, \cdots, cx_{n+1})$确定 P 的同一个点，所以 T^* 的确

是恒等变换．反过来，假设 T^* 是恒等变换，那么 T 一定把 $n+1$ 个单位矢量 $\varepsilon_1, \cdots, \varepsilon_{n+1}$ 的每一个 ε_i 变换到某个数乘积 $c_i\varepsilon_i$，因此 A 一定是对角矩阵，其对角线元素为 $c_1, \cdots, c_{n+1}$．但是 T 也把矢量 $(1, 1, \cdots, 1)$ 变换到它的一个数乘积，而 A 把这个矢量变换到 $(c_1, \cdots, c_{n+1})$．$(c_1, \cdots, c_{n+1})$ 是 $(1, 1, \cdots, 1)$ 的数乘积当且仅当 $c_1, \cdots, c_{n+1}$ 都相等．因此 A 的确是 I 的一个数乘积．

推论 域 F 上的 n 维射影群与 $n+1$ 维全线性群对于由恒等变换的非零数乘积组成的子群的商群同构．

证明 映射 $T \longmapsto T^*$ 是全线性群到射影群中的同态；定理 33 断言，这个同态的核恰恰是由恒等变换的数乘积组成的集合．因此根据 § 7.13 定理 28 得此结论．

还可以推出，两个矩阵 A 和 A_1 确定同一个射影变换当且仅当 $A_1 = cA$（c 为某一标量）．

对于一维射影直线，射影变换具有形式

$$y_1 = ax_1 + bx_2, \quad y_2 = cx_1 + dx_2, \quad ad \neq bc. \tag{61}$$

按照非齐次坐标 $z = \frac{x_1}{x_2}$，$w = \frac{y_1}{y_2}$，这个变换可写成线性分式变换

$$w = \frac{az+b}{cz+d}, \tag{62}$$

它是由(61)的第一个方程除以第二个方程而得到的．公式(62)可以解释如下．如果 $c=0$，则(62)把点 $z=\infty$ 变到点 $w=\infty$；如果 $c \neq 0$，则(62)把点 $z=\infty$ 变到点 $w=\frac{a}{c}$，把点 $z=-\frac{d}{c}$ 变到点 $w=\infty$．这些解释的正确性可以通过代回齐次坐标并利用(61)式来验证．在 n 维的情况，也可能有用线性分式变换表示射影变换的类似的表达式

$$w_i = \frac{z_1 a_{1i} + \cdots + z_n a_{ni} + a_{n+1,i}}{z_1 b_1 + \cdots + z_n b_n + b_{n+1}}, \quad (b_j = a_{j,n+1}; \ i = 1, \cdots, n). \tag{62'}$$

我们已经看到，$P_n(F)$的射影变换把直线变到直线．反过来，实射影空间 $P_n(\mathrm{R})$ 的任意把直线变到直线的一一变换，当 $n\geqslant 2$ 时，是射影变换(见习题 6)，这是一个经典的结果．

三个变量的二次型确定射影平面中的一个轨迹

$$\sum_{i,j} x_i b_{ij} x_j = 0 \quad (i, j = 1, 2, 3), \tag{63}$$

这是因为如果坐标 (x_1, x_2, x_3) 满足这个方程，那么任意"数乘"积 (cx_1, cx_2, cx_3)也满足这个方程．这个轨迹称为射影二次曲线；这个二次曲线的(射影)秩是系数矩阵B的秩．如果排除"在无穷远处的直线"，这个射影二次曲线(63)就变成普通二次曲线．在实射影平面中，任意非退化二次曲线(即椭圆，双曲线或抛物线)，根据§ 9.9，它与下列四个方程中的一个所确定的曲线等价：

$$x_1^2 + x_2^2 + x_3^2 = 0, \qquad x_1^2 + x_2^2 - x_3^2 = 0, \tag{64}$$

$$-x_1^2 - x_2^2 - x_3^2 = 0, \quad x_1^2 - x_2^2 - x_3^2 = 0. \tag{64'}$$

这些方程中，等号左边各项符号改变并不改变其轨迹，因此由(64′)给出的二次曲线本质上就是(64)给出的二次曲线．(64)第一条二次曲线是空的．因此我们得出，任意两个非退化二次曲线在实射影平面中是射影等价的．

习　　题

1．在域F上的三维射影空间中，证明：

(a) 任意两个不同的点在一条且只在一条直线上．

(b) 任意不在一直线的三个点在一个且只在一个平面上．

2．推广习题 1 到 n 维射影空间．

3．在域 $\mathbf{Z}_2$ 上的射影平面中列出所有的点和直线，以及每条直线上的全部点．

4．在含有 n 个元素的有限域上的射影平面中，证明：存在 n^2+n+1 个

点，n^2+n+1 条直线，并且每条直线上有 $n+1$ 个点.

5. 四个不同数 z_1, z_2, z_3, z_4 的交比定义为比值 $\frac{(z_3-z_1)(z_4-z_2)}{(z_3-z_2)(z_4-z_1)}$（当 z_i 中有一个是∞时，用适当的约定）. 证明：交比是任意线性分式变换 (62) 之下的不变量.

6. 证明：在复射影平面上，变换 $(z_1, z_2, z_3)\mapsto(z_1^*, z_2^*, z_3^*)$ 把直线变到直线，但这个变换不是射影变换.（星号表示复共轭.）

7. 如果去掉"在无穷远外的直线"$x_3=0$，那么射影二次曲线 $x_1^2=2x_2x_3$ 在仿射平面中表示什么？

8. (a) 证明：每个非退化实二次曲面与一个球面或与一个单叶双曲面射影等价.

(b) 椭圆抛物面与上面哪个曲面射影等价？双曲抛物面与上面哪个曲面射影等价？

(c) 证明：球面与单叶双曲面不射影等价.

9. 证明：在射影直线中给出任意两个不同点的 3-数组 (z_1, z_2, z_3) 和 (w_1, w_2, w_3)，则存在一个射影变换 (62)，它把每个 z_i 变到相应的 w_i.

10. 设 (p_1, p_2, p_3, p_4) 和 (q_1, q_2, q_3, q_4) 是射影平面中的任意两个点的 4-数组. 证明：存在一个射影变换 (62′)，它把每个 p_i 变到相应的 q_i.

第十章　行列式与标准型

§10.1　行列式的定义和基本性质

在任意域上每个方阵 A 都有一个行列式；虽然行列式能够用来研究矩阵的秩和求解联立线性方程组，但是它在矩阵论中最重要的应用是定义矩阵的特征多项式．在这一章中，我们来定义行列式，研究它的几何性质，并指出矩阵 A 的特征多项式和矩阵的特征根(特征值)之间的关系．然后，用这些概念来研究矩阵在相似变换之下的标准型．

联立线性方程组的求解公式自然导出行列式．两个线性方程

$$a_1x+b_1y=k_1,$$
$$a_2x+b_2y=k_2$$

在 $a_1b_2-a_2b_1\neq 0$ 的假定下，有唯一解

$$x=\frac{k_1b_2-k_2b_1}{a_1b_2-a_2b_1},\quad y=\frac{a_1k_2-a_2k_1}{a_1b_2-a_2b_1},$$

出现在分子和分母中的多项式称为行列式，

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}=a_1b_2-a_2b_1,\qquad \begin{vmatrix} k_1 & b_1 \\ k_2 & b_2 \end{vmatrix}=k_1b_2-k_2b_1. \tag{1}$$

类似地，我们可以计算三个联立线性方程组 $\sum a_{ij}x_j=k_i$ 的解．结果每个解 x_j 的分母是

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}=\begin{aligned}&a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}\\ &-a_{11}a_{23}a_{32}-a_{12}a_{21}a_{33}-a_{13}a_{22}a_{31},\end{aligned} \tag{2}$$

等式的右边是六项乘积的代数和．每一项包含矩阵第一行的某元

素 a_{1i}，同时也包含第二行的某元素和第三行的某元素．每一项还可以表示成包含着不同列的元素乘积，所以(2)中的每一项可写成 $a_{1-}a_{2-}a_{3-}$，空白的地方添上列标 1，2，3 的某一置换．在六个可能的置换中，三个偶置换 I，(123)，(132)出现在带正号的乘积项中，而三个奇置换出现在带负号的乘积项中．经验表明，n 个未知数的 n 个线性方程的解可以表示为类似的公式．

定义 $n\times n$ 矩阵 $A=(a_{ij})$ 的行列式是元素[①] $a_{ij}=a(i,j)$ 的多项式：

$$\det A=|A|=\sum_{\phi}\operatorname{sgn}\phi\left[\prod_{i=1}^{n}a_{i,i\phi}\right]$$

$$=\sum_{\phi}(\operatorname{sgn}\phi)a(1,1\phi)a(2,2\phi)\cdots a(n,n\phi). \tag{3}$$

这里是对整数 1，…，n 的所有 $n!$ 个不同的置换 ϕ 求和．每个乘积项 $\Pi a_{i,i\phi}$ 前面的因子 $\operatorname{sgn}\phi$ 是 $+1$ 还是 -1，根据 ϕ 是偶置换还是奇置换而定．

于是行列式 $|A|=\det(a_{ij})$ 是 $n!$ 项 $\pm a_{1-}a_{2-}\cdots a_{n-}$ 的和，这里空白处添上数字 1，…，n 的各种置换 ϕ．把 a_{ij} 写成 $a(i,j)$，设 $i\phi$ 是 i 在 ϕ 之下的象，于是一般项可写成 $\pm a(1,1\phi)a(2,2\phi)\cdots a(n,n\phi)$，这里正负号“±”称为 $\operatorname{sgn}\phi$（signumϕ 的缩写）．对于和中的每个乘积项，恰好含有矩阵每一行中一个元素，并且恰好含有矩阵每一列中一个元素．

矩阵的每一行在 $|A|$ 的每一项中出现一次且只出现一次，这就意味着 $|A|$ 是 A 的第 i 行元素 $a_{i1},\cdots,a_{in}$ 的线性齐次函数．把每个这样的 a_{ij} 的系数合并起来，我们就得到表达式

$$|A|=A_{i1}a_{i1}+A_{i2}a_{i2}+\cdots+A_{in}a_{in}, \tag{4}$$

这里 a_{ij} 的系数 A_{ij} 称为 a_{ij} 的余子式（即代数余子式）；它是 A 中

① 这里指的元素是域 F 中的元素，或更一般地是指交换环中的元素．

划去第 i 行剩下的各行元素的多项式. 这个余子式还可以描述成偏导数 $A_{ij}=\frac{\partial|A|}{\partial a_{ij}}$. 因为 $|A|$ 的每一项只包含每一行和每一列中一个元素, 所以余子式 A_{ij} 既不包含第 i 行的元素也不包含第 j 列的元素. 它只包含"子式"或子矩阵 M_{ij} 的元素, 子矩阵 M_{ij} 是从矩阵 A 中划去第 i 行和第 j 列的所有元素所得到的矩阵.

$|A|$ 对于行和列是对称的.

定理 1 设 A^{T} 表示 A 的转置矩阵, 则 $|A^{T}|=|A|$.

证明 A^{T} 的元素 $a_{ij}^{T}=a_{ji}$ 是把 A 的元素 a_{ij} 的下标颠倒过来而得到的. 当 $j=i\phi, i=j\phi^{-1}$, $|A|$ 的一般项是

$$(\mathrm{sgn}\phi)\prod_{i}a(i,i\phi)=(\mathrm{sgn}\phi)\prod_{j}a(j\phi^{-1},j)$$

$$=(\mathrm{sgn}\phi)\prod_{j}a^{T}(j,j\phi^{-1}),$$

它也是 $|A^{T}|$ 的一般项, 因为每一置换是某个置换 ϕ 的逆置换 ϕ^{-1}, 而且正负号也相同, $\mathrm{sgn}\phi=\mathrm{sgn}\phi^{-1}$, 这是根据 ϕ 是偶置换(即在交错群中) 当且仅当 ϕ 的逆 ϕ^{-1} 也是偶置换 (§ 6.10). 因此 $|A|=|A^{T}|$. 证毕

对行列式进行初等行运算时, 将会产生什么样的效果呢?

法则 1 矩阵 A 的第 i 行乘上标量 $c\neq0$, 则相应的行列式就乘上 c. 这是因为在线性齐次表达式 (4) 中, 对第 i 行的每个元素 $a_{i1},\cdots,a_{in}$ 乘上一个因子 c 就意味着对 $|A|$ 乘上同一个因子 c.

法则 2 矩阵 A 的两行交换, 行列式 $|A|$ 变号. 根据对称性(定理 1), 我们可以改为证明矩阵的两列交换, 行列式变号. 这种交换可以用列标的奇置换 ϕ_0 来表示, 于是可用矩阵 $B=(b_{ij})$ 来代替 A, 其中 $b_{ij}=b(i,j)=a(i,j\phi_0)$, 那么

$$|B|=\sum_{\phi}(\mathrm{sgn}\phi)\prod_{i}b(i,i\phi)=\sum_{\phi}(\mathrm{sgn}\phi)\prod_{i}a(i,i\phi\phi_0)$$

因为全体置换构成一个群, 所以全体乘积 $\phi\phi_0$(这里 ϕ_0 是固定的,

ϕ 跑遍所有置换）包含全部置换，因此上面的$|B|$含有$|A|$的所有项．只是每一项的正负号改变了，这是因为ϕ_0是奇置换，所以当ϕ是奇置换时，$\phi\phi_0$是偶置换，反过来，当ϕ是偶置换时，$\phi\phi_0$是奇置换，于是 $\operatorname{sgn}\phi\phi_0=-\operatorname{sgn}\phi$．这就证明了法则 2．

引理 1 如果矩阵A有相同的两行，那么$|A|=0$．

证明 根据定理 1，只须证明当A有相同的两列时，$|A|=0$．设ψ是把两个相同的列进行交换的对换，那么表达式(3)中的所有被加项$(\operatorname{sgn}\phi)\Pi a(i,i\phi)$是按$\{\phi,\psi\phi\}$成对出现，而$\{\phi,\psi\phi\}$是由$\psi$生成的二元素子群的陪集组成．因为$\psi$是奇置换，所以$\operatorname{sgn}\phi=-\operatorname{sgn}\psi\phi$；又因为两个列是相同的，所以有$\Pi a(i,i\phi)=\Pi a(i,i\psi\phi)$．因此成对的被加项其数值相等符号相反，所以它们的和是零． 证毕

为了考虑伴随矩阵（§ 10.2），用一个方程表示引理 1 是方便的．在A中，用第k行来代替第i行，则两行变成一样了，其行列式等于零．但是这个行列式还可以在线性齐次表达式（4）中用第k行代替第i行而得到，于是

$$0=A_{i1}a_{k1}+A_{i2}a_{k2}+\cdots+A_{in}a_{kn}\quad(i\neq k). \tag{5}$$

法则 3 矩阵第k行乘上常数c加到第i行上，其行列式保持不变．这个运算是用$a_{ij}+ca_{kj}$代替每个a_{ij}；根据线性齐次表达式(4)，新的行列式是

$$\sum_j A_{ij}(a_{ij}+ca_{kj})=\sum_j A_{ij}a_{ij}+c\sum_j A_{ij}a_{kj}=|A|+0,$$

上式最后一个等式是根据(4)，(5)得到的．行列式的确没变．

这些法则可以用初等矩阵加以概括．任意初等行运算把单位矩阵I变成初等矩阵E，而把A变成乘积EA．行列式$|I|=1$因而变成$|E|=c$，-1或 1（对应于法则 1，2 或 3），而$|A|$变成$|EA|=c|A|$，$-|A|$或$|A|$（对应于相应的法则）．这就证明了$|EA|=$

$|E||A|$. 根据对称性(定理 1), 同样可应用于右乘因子 E. 这就建立了

定理 2 *如果 E 是初等矩阵, 那么*

$$|EA|=|E||A|=|AE|.$$

另一个法则是, 从上面讨论的子矩阵 M_{ij} 可以明显地得到余子式的表达式.

法则 4 $A_{ij}=(-1)^{i+j}|M_{ij}|$,口头上说就是,每个余子式$A_{ij}$是由相应的子矩阵的行列式前面添上符号$(-1)^{i+j}$而得到. 正负号$(-1)^{i+j}$可以在西洋跳棋盘 (想象在方格盘中 +, − 号相间排列, 并且左上角第一格中是 + 号)上的(i, j)位置上得到. 首先我们对 $i=j=1$ 来证明这个法则. 从定义(3)立即看出, 包含 a_{11} 的项恰恰是满足条件 $1\phi=1$ 的那些置换 ϕ 所对应的项. 这种类型的偶置换(或奇置换) 实际上是剩下的数字 $2, \cdots, n$ 的偶置换 (或奇置换),所以这些项划去 a_{11} 后恰恰就是$|M_{11}|$的展开式的所有项. 任意其他余子式 A_{ij}, 可通过把 a_{ij} 移到左上角而化成上述特殊情形, 而 a_{ij} 移到左上角位置是相继进行 $i-1$ 次相邻两行的交换和 $j-1$ 次相邻两列的交换得到的. 这些初等运算并不改变$|M_{ij}|$, 因为它不影响 M_{ij} 中行与列的相互位置, 但是这些运算却改变了$|A|$的符号, 因此 a_{ij} 的余子式的符号改变了 $i+j-1-1$ 次. 通过这样的简化就证明了法则 4.

一个特别有用的情形是, 矩阵的第一行除了第一个元素外其他所有元素都是零. 那么表达式(4)只须包含第一个余子式$|M_{11}|=A_{11}$,所以

$$\begin{vmatrix} c & O \\ K & B \end{vmatrix}=c|B|, \tag{6}$$

这里 O 是$1\times(n-1)$矩阵, K 是 $(n-1)\times 1$ 矩阵, B 是$(n-1)\times(n-1)$矩阵. 根据这个法则并用归纳法, 我们得到下面结果.

引理 2 三角形矩阵的行列式等于它的对角线元素之积.

上述各法则提供了一系列计算行列式 $|A|$ 的方法. 通过初等运算把矩阵 A 化为三角形矩阵 T, 用 t 表示所用行(或列)交换的次数, 用 $c_1, \cdots, c_s$ 表示乘到 A 的行(或列)上的各个标量. 根据定理2, $|A|=(-1)^t(c_1\cdots c_s)^{-1}|T|$, 再用引理 2 设 $|T|=t_{11}\cdots t_{nn}$, 这样就计算出行列式的值.

习 题

1. 从行列式的定义直接证明引理 2.

2. 计算 §7.6 习题 2 中各矩阵的行列式.

3. (a) 设 $A=\begin{pmatrix} 1 & -1 & 0 \\ -1 & 0 & 1 \\ 2 & 1 & 1 \end{pmatrix}$. 分别用第一行的子式或第一列的子式计算行列式 $|A|$, 并比较这两个计算结果.

(b) 假定上面矩阵 A 的元素是模 2 整数, 计算 $|A|$.

4. 写出一般 4×4 行列式展开式中的所有正项.

5. 设 n 是奇数, 并且 $1+1\neq 0$, 证明: $n\times n$ 斜对称矩阵 A 的行列式为零.

6. (a) 推导下面范得蒙(Vandermonde)行列式的展开式

$$\begin{vmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \end{vmatrix}=(x_2-x_1)(x_3-x_1)(x_3-x_2).$$

(b) 推广这个结果到 4×4 的情形.

(c) 推广这个结果到 $n\times n$ 的情形, 即证明: 如果 $a_{ij}=x_i^{j-1}$, 那么 $|A|=\prod_{i>j}(x_i-x_j)$.

7. 证明: 对任意 4×4 斜对称矩阵 A, 有 $|A|=(a_{12}a_{34}-a_{13}a_{24}-a_{14}a_{23})^2$.

8. (a) 证明: 任意置换矩阵的行列式等于 ± 1.

(b) 证明: 单项矩阵的行列式等于全体非零元素的乘积再乘上 ± 1.

9. 如果 $n\times n$ 实矩阵 A, 对于 $i=1, \cdots, n$, 有 $\sum_{j\neq i}|a_{ij}|<a_{ii}$, 则称 A 为对

角优势矩阵. 证明: 如果A是对角优势矩阵, 那么$|A|>0$.

10. 证明: 平面上两个点(a_1, a_2)和(b_1, b_2)联线的方程是

$$\begin{vmatrix} x_1 & x_2 & 1 \\ a_1 & a_2 & 1 \\ b_1 & b_2 & 1 \end{vmatrix}=0.$$

*11. (a) 证明: 如果矩阵A的每个元素a_{ij}是x的函数, 那么$\frac{d|A|}{dx}=\sum_{j,k=1}^{n}\frac{da_{jk}}{dx}A_{jk}$.

(b) 用此验证$A_{ij}=\frac{\partial|A|}{\partial a_{ij}}$.

*12. 证明: 如果A和C是方阵, 那么有

$$\begin{vmatrix} A & B \\ O & C \end{vmatrix}=|A|\,|C|$$

*13. 设Ω是$n\times n$矩阵(ω^{ij}), 这里ω是n次本原复单位根, 证明: 如果$n\equiv1(\mathrm{mod}4)$, 那么有$|\Omega|=n^{\frac{n}{2}}$.

§10.2 行列式的乘积

在初等行运算与初等列运算之下, 任意方阵A等价于对角矩阵D(§8.9定理18), 所以象§8.8定理13那样, 矩阵A可以由矩阵D左乘和右乘一些初等矩阵E_i和$E^{(i)}$得到, 即

$$A=E_s\cdots E_1DE^{(1)}\cdots E^{(t)} \tag{7}$$

根据定理2的法则$|EA|=|E|\,|A|$和$|AE|=|A|\,|E|$, 在乘积(7)的行列式中, 可以同时对每个因子E取行列式$|E|$, 于是得到

$$|A|=|E_s|\cdots|E_1|\,|D|\,|E^{(1)}|\cdots|E^{(t)}|. \tag{8}$$

因为每个$|E|\neq0$, 所以整个行列式$|A|\neq0$当且仅当$|D|\neq0$. 这个标准型D在主对角线上恰恰有r个元素1, 这里r是矩阵A的秩, 而行列式$|D|$是它的n个对角线元素之积. 因此$|D|\neq0$当且仅当$r=n$, 也就是当且仅当A是非奇异的. 所以由(8)式证明了

定理 3　方阵 A 是非奇异的当且仅当 $|A|\neq 0$.

行列式的计算　公式(8)也为 $n\times n$ 行列式的数值计算提供了一个有效方法．我们象高斯消去法那样进行计算，逐个把对角线元素用 1 代替，这样容易算出对角线元素的乘积；因为除了某一行(或列)乘以常数的初等矩阵外，其他所用的初等矩阵的行列式都是 ± 1，所以这足以计算出整个行列式．例如，

$$\begin{vmatrix} 2 & 3 & 4 & 1 \\ 4 & -1 & 2 & 3 \\ -6 & 5 & 2 & 6 \\ 8 & 5 & 7 & -2 \end{vmatrix} = 2\begin{vmatrix} 1 & \frac{3}{2} & 2 & \frac{1}{2} \\ 0 & -7 & -6 & 1 \\ 0 & 14 & 14 & 9 \\ 0 & -7 & -9 & -6 \end{vmatrix}$$

$$= -14\begin{vmatrix} 1 & \frac{3}{2} & 2 & \frac{1}{2} \\ 0 & 1 & \frac{6}{7} & -\frac{1}{7} \\ 0 & 0 & 2 & 11 \\ 0 & 0 & -3 & -7 \end{vmatrix}$$

因此这个行列式等于 $(-14)(19)=-266$.

一个非奇异矩阵 A 是初等矩阵的乘积 $A=E_t\cdots E_1$．如果 $B=E_s^*\cdots E_1^*$ 是另一个这样的矩阵，那么乘积 AB 的行列式可象(8)式那样进行计算，结果为

$$\begin{aligned}|AB| &= |E_t\cdots E_1E_s^*\cdots E_1^*| \\ &= |E_t|\cdots|E_1|\cdot|E_s^*|\cdots|E_1^*| = |A|\cdot|B|\end{aligned}$$

定理 4　乘积矩阵的行列式是各矩阵的行列式的乘积：$|AB|=|A|\cdot|B|$.

证明　上述计算仅给出当 A 和 B 都是非奇异矩阵时，这个法则的证明．而当 A 或 B 是奇异矩阵时，因此 AB 也是奇异矩阵，所以 $|AB|=|A|\cdot|B|$ 的两边都等于零。　　证毕

行列式$|A|\neq 0$的矩阵A的逆存在，并且可以用A的余子式明显地求出．最初的包含余子式的方程(4)和(5)可以写成

$$a_{k1}A_{i1}+\cdots+a_{kn}A_{in}=\delta_{ki}|A|,\qquad \delta_{ki}=\begin{cases}1, & \text{当 } i=k,\\ 0, & \text{当 } i\neq k.\end{cases}\tag{9}$$

δ_{ki}这个数恰好是单位矩阵$I=(\delta_{ki})$的(k, i)位置上的元素．方程(9)很象矩阵的乘积，如果把余子式的下标交换，那么(9)式左边就给出了矩阵A和余子式矩阵的转置矩阵的乘积的(k, i)位置上的元素．(9)式右边是用标量$|A|$乘上单位矩阵后的(k, i)位置上的元素，所以有

$$A(A_{ij})^{\tau}=|A|I.\tag{10}$$

出现在这个方程中的矩阵$(A_{ij})^{\tau}$是A的全体元素的余子式构成的矩阵的转置矩阵，称为A的*伴随矩阵*．在$|A|=1$的情况下，方程(10)表明，A的伴随矩阵就是A的逆矩阵．一般地，如果$|A|\neq 0$，(10)式证明了

定理 5 如果$|A|\neq 0$，那么矩阵A的逆是$A^{-1}=|A|^{-1}(A_{ij})^{\tau}$.

解n个未知数n个线性方程的克莱姆法则是这个逆矩阵公式的推论．已知的线性方程组形为

$$\sum_j a_{ij}x_j=b_i,$$

其中i和j是从1到n．用矩阵表示，这个方程组就是$AX=B$(X和$B=(b_1, \cdots, b_n)^{\tau}$都是$n$维列矢量)．如果$A$是非奇异的，那么用$A^{-1}$左乘这个方程得到唯一的矢量解$X=(x_1, \cdots, x_n)^{\tau}=A^{-1}B$．如果我们注意到逆矩阵$A^{-1}$的$(i, j)$位置上的元素刚好是$\frac{A_{ji}}{|A|}$，那么这个解可以展开表出．这就证明了

定理 6 (克莱姆法则) 如果n个未知数n个线性方程

$$\sum_j a_{ij}x_j=b_i$$

的系数矩阵 $A=(a_{ij})$ 是非奇异的，那么方程组有唯一解

$$x_j=\frac{A_{1j}b_1+\cdots+A_{nj}b_n}{|A|}, \quad j=1, \cdots, n, \tag{11}$$

其中 A_{ij} 是系数矩阵 A 中元素 a_{ij} 的余子式。

这个公式的分子本身可以写成行列式，因为它是用常数列 $(b_1, \cdots, b_n)'$ 代替矩阵 A 的第 j 列而得到的行列式按第 j 列的余子式的展开式。可是对于大的联立线性方程组，通过化矩阵（或增广矩阵）为行等价梯矩阵（§7.7）的方法求解，通常更为有效。

显然，克莱姆法则可以应用到任意域上，特别可以应用到§2.3中所讨论的所有方程组（参看下面习题9）。克莱姆法则对于求解二个或三个未知数的联立线性方程组是特别方便的。

附录　行列式与秩　长方矩阵 A 的子矩阵（或"子式"）是从 A 划去 A 的某些行和某些列而得到的任意矩阵（这里包括一行也没划去或一列也没划去的情况）。任意长方矩阵 $A\neq O$ 的"行列式秩" d 可以定义为行列式不为零的 A 的最大子式的行数，换句话说，d 具有性质：(i) A 至少有一个 $d\times d$ 子式 M，它的行列式 $|M|\neq 0$；(ii) 如果 $h>d$，则 A 的每个 $h\times h$ 子式 N，都有 $|N|=0$。可以证明，任意矩阵的秩等于它的行列式秩。

习　　题

1. 写出 2×2 矩阵 $A=\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ 的伴随矩阵，并计算 A 和它的伴随矩阵的乘积。

2. (a) 计算§7.6习题2(a)中的矩阵的伴随矩阵。对于这种情形验证关于矩阵和它的伴随矩阵的乘积法则。

   (b) 对§7.6习题2(b)做同样的问题。

3. 用伴随矩阵的方法，求§8.8(50)式中的 4×4 初等矩阵 H_{24}, $I+2E_{33}$ 和 $I+dE_{21}$ 的逆矩阵。

4. 用伴随矩阵的方法，求§8.8习题5的矩阵的逆.

5. 证明：如果A是非奇异矩阵，那么$|A^{-1}|=|A|^{-1}$.

6. 证明：奇异矩阵和它的伴随矩阵的乘积是零矩阵.

7. 证明：任意正交矩阵的伴随矩阵是它的转置矩阵.

8. 写出三个未知数三个线性方程的克莱姆法则.

9. 用克莱姆法则，求解§2.3习题1的联立同余式.

10. (a) 证明：下面一对齐次线性方程

$$a_1x+b_1y+c_1z=0,$$

$$a_2x+b_2y+c_2z=0$$

有一组解

$$x=\begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix},\quad y=\begin{vmatrix} c_1 & a_1 \\ c_2 & a_2 \end{vmatrix},\quad z=\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}.$$

(b) 什么时候这组解是整个解集合的基底?

(c) 对四个未知数的三个线性方程推导类似的公式.

11. 证明：正交矩阵的行列式是±1.

12. 证明：矩阵A的伴随矩阵的行列式是$|A|^{n-1}$.

13. 证明：矩阵A的伴随矩阵的伴随矩阵是$|A|^{n-2}A$.

14. 直接由行列式秩的定义证明：初等行运算不改变行列式秩.

*15. (a) 设A和B是3×3矩阵，证明：AB的任意2×2子矩阵的行列式是一些项的和，其中每一项是A的一个2×2子矩阵的行列式与B的一个2×2子矩阵的行列式的乘积.

(b) 推广这个结果，并用它证明$\operatorname{rank}(AB)\leqslant\operatorname{rank}A$.

*16. 设$n\times n$矩阵A的秩为r，证明：A的伴随矩阵的秩s确定如下：若$r=n$，则$s=n$；若$r=n-1$，则$s=1$；若$r<n-1$，则$s=0$.

*17. 证明：任意矩阵的秩等于它的行列式秩.

§10.3 作为体积的行列式

$n\times n$实矩阵的行列式在几何上可以解释为n维欧几里得空间中的体积. 这是从平行四边形面积公式得到启发的.

每个以矢量α_1和α_2为行的2×2实矩阵A表示一个以O, α_1, α_2, $\alpha_1+\alpha_2$为顶点的平行四边形. 反过来，每个这样的平行四边形

确定一个矩阵(参看图 1). 这个平行四边形的面积是

$$底\times高=|\alpha_1||\alpha_2||\sin C|, \tag{12}$$

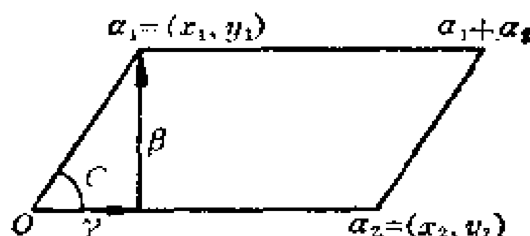


图 1

式中C表示已知矢量α_1和α_2之间的夹角. 根据 § 7.9 的余弦公式(41), 这个面积的平方等于

$$(\alpha_1,\alpha_1)(\alpha_2,\alpha_2)(1-\cos^2C)=(\alpha_1,\alpha_1)(\alpha_2,\alpha_2)-(\alpha_1,\alpha_2)(\alpha_2,\alpha_1).$$

这个结果看起来很象 2×2 矩阵的行列式, 实际上, 它就是矩阵 $((\alpha_i,\alpha_j))=AA^T$ 的行列式.

任意维欧几里得空间中的平行四边形都具有类似的公式, 甚至还可以推广到n维欧几里得空间中的m维类似于平行四边形的几何体, 这些几何体称为*平行六面体*.

为了进行这种推广, 设A是任意一个以$\alpha_1,\cdots,\alpha_m$为行的$m\times n$矩阵. 这些行矢量表示在n维欧几里得空间E_n中从原点出发的矢量. E_n中由m个矢量$\alpha_1,\cdots,\alpha_m$张成的平行六面体Π是由所有形为

$$t_1\alpha_1+\cdots+t_m\alpha_m \quad (0\leqslant t_i\leqslant 1;\ i=1,\cdots,m)$$

的矢量组成. (对 $m=n=3$ 的情形画出图形, 你会得到一个与立方体仿射等价的平行六面体!) 这样就建立起 $m\times n$ 实矩阵与n维空间中的m维平行六面体之间的对应, $\alpha_1,\cdots,\alpha_m$称为这个平行六面体Π的棱.

这个平行六面体的m维体积$V(\Pi)$(包括当$m=1$时的长度和当$m=2$时的面积这两种特殊情形)可以对m用归纳法来定义. 设以$\alpha_2,\cdots,\alpha_m$为棱的平行六面体称为Π的底. α_1的与$\alpha_2,\cdots,\alpha_m$正

交的分量称为高，它是通过把剩下的一条棱α_1写成两个分量γ与β的和来求得，其中分量γ在$\alpha_2, \cdots, \alpha_m$张成的空间$S_{m-1}$中，$\beta$与$S_{m-1}$正交(见图1，根据§7.11，这总是可能的)：

$$\alpha_1=\beta+\gamma, \quad \beta\perp S_{m-1}, \quad \gamma \text{ 在 } S_{m-1} \text{ 中}. \tag{13}$$

Π的体积就定义为底的$m-1$维体积与高的长度$|\beta|$的乘积.

定理 7 *以$\alpha_1, \cdots, \alpha_m$为棱的平行六面体的体积的平方是行列式$|AA^T|$，其中$A$是以矢量$\alpha_i$的各坐标*①*为第$i$行元素的矩阵.*

注　因为A的行置换是用PA代替A，其中P是满足$|P|=|P^T|=\pm1$的$m\times m$置换矩阵，又因为

$$|(PA)(PA)^T|=|P|\cdot|AA^T|\cdot|P^T|=|AA^T|,$$

所以Π的“体积”与它的底是由哪$m-1$个矢量张成的无关.

证明　因为A是$m\times n$矩阵，所以乘积AA^T是一个$m\times m$方阵. 现在我们对m归纳进行论证. 当$m=1$时，矩阵A是行矩阵，“内积”$AA^T=(\alpha_1, \alpha_1)$是$\alpha_1$的长度的平方，满足要求. 假定对于$m-1$行的矩阵，这个定理是正确的，我们来考虑$m$行的情形. 象(13)式那样，第一行$A_1$可以写成$A_1=B_1+C_1$，这里“高”$B_1$同每个行矢量$A_2, \cdots, A_m$都正交(即$B_1A_i^T=0, i=2, \cdots, m$)，而$C_1=c_2A_2+\cdots+c_mA_m$是$A_2, \cdots, A_m$的线性组合. 从$A$的第一行逐次减去第$i$行的$c_i$倍($i=2, \cdots, m$)，这就把$A$变成新的矩阵$A^*$，它的第一行是$B_1$. 而且，这些初等行运算每一个都相当于用一个行列式为1的初等矩阵左乘矩阵A，因此$A^*=PA$，这里$|P|=1$，并且有$|A^*A^{*T}|=|PAA^TP^T|=|P||AA^T||P^T|=|AA^T|$. 如果$D$是由$A^*$的$m-1$行$A_2, \cdots, A_m$组成的矩阵块，那么

$$A^*A^{*T}=\begin{pmatrix}B_1\\D\end{pmatrix}(B_1^T\ D^T)=\begin{pmatrix}B_1B_1^T & B_1D^T\\DB_1^T & DD^T\end{pmatrix}$$

---

① 整个§10.3中，矢量的坐标都是对于一组固定的标准正交基来取的. 当$m>n$时，定理7退化为方程$0=0$.

$$=\begin{pmatrix} B_1B_1^{\mathrm{T}} & O \\ O & DD^{\mathrm{T}} \end{pmatrix}$$

这里 $B_1D^{\mathrm{T}}=O$ 是因为对 D 的每行 A_i 有 $B_1A_i^{\mathrm{T}}=0$. 根据(6)式, 行列式就是

$$|AA^{\mathrm{T}}|=|A^*A^{*\mathrm{T}}|=(B_1B_1^{\mathrm{T}})|DD^{\mathrm{T}}|.$$

这里 D 是矩阵, 它的所有行 $A_2, \cdots, A_m$ 张成 Π 的底, 所以根据归纳法假设, $|DD^{\mathrm{T}}|$ 是底的体积的平方. 此外, 标量 B_1B^{T} 是高的长度的平方, 所以我们得到所要求的关于 AA^{T} 的底×高的公式.

证毕

在行数为 n 的特殊情形中, 显然有 $|AA^{\mathrm{T}}|=|A|\cdot|A^{\mathrm{T}}|=|A|^2$, 于是我们就证明了①

定理 8 设 A 是任意 $n\times n$ 实矩阵, 它的各行为 $\alpha_1, \cdots, \alpha_n$, 则 A 的行列式(除了可能相差一个正负号外)是 E_n 中以 $\alpha_1, \cdots, \alpha_n$ 为棱的平行六面体的体积.

行的任意置换并不改变行列式的绝对值, 所以这个定理还表明, 平行六面体的体积定义与棱的排列顺序无关. 当 $m<n$ 时, 这个论证还可用到定理 7 的公式. 当 $m=n$ 时, 行列式 $|A|$ 常常称为以 $\alpha_1, \cdots, \alpha_n$ 为棱的平行六面体的"带符号"的体积. 用任意奇置换可使它的符号改变.

定理 9 一个 n 维欧几里得矢量空间的线性变换 $Y=XP$, 使得所有 n 维平行六面体的体积乘上因子 $\pm|P|$.

证明 考虑一个平行六面体, 它分别以行矢量 $A_1, \cdots, A_n$ 为 n 个棱. 行矢量 $A_1, \cdots, A_n$ 被变成 $A_1P, \cdots, A_nP$: 含有这些新行矢量的矩阵就是矩阵乘积 AP, 其中 A 具有行 $A_1, \cdots, A_n$. 那么新的"带符号"的体积等于 $|AP|=|A||P|$, 这里 $|A|$ 是原来平行六面体

① 证明这个定理的基本步骤最初是由 J. S. Frame 教授提出的.

的体积.

由此推出：变换 $Y=XP$ 使"带符号"的体积保持不变当且仅当变换矩阵满足 $|P|=1$. 所有具有这种性质的矩阵（或所有变换）组成的集合称为幺模群. 有时这种群扩大成包含所有满足 $|P|=\pm1$ 的矩阵（也就是，保持体积绝对值不变的所有变换）.

n 维欧几里得空间中，任意区域 f 的体积可以粗略地定义如下：用一组有限个给定形状和方位的平行六面体 $\Pi_1, \cdots, \Pi_s$ 外接于 f，然后求和 $\sum V(\Pi_i)$，定义 f 的体积是所有这种平行六面体的不同集合的体积和的最大下界（第四章）.（在积分学中，这一般是可以做到的，那些平行六面体是其边平行于坐标轴的立方体.）

根据定理 9，具有矩阵 P 的线性变换以 $1:|P|$ 的比例改变任意平行六面体的体积，因此它也以同样的比例改变 f 的体积. 因为平移使体积保持不变，所以我们得到下面结果.

推论 仿射变换 $Y=XP+K$ 使所有体积乘上一个因子 $|P|$（或更确切地说，是乘上 $|P|$ 的绝对值）.

习　　题

1. (a) 计算平面上以 $(0, 0)$, $(3, 0)$, $(1, 4)$ 和 $(4, 4)$ 为顶点的平行四边形的面积.

   (b) 计算空间中以 $(0, 2, 0)$, $(2, 0, 0)$, $(1, 1, 5)$ 和 $(0, 0, 0)$ 为相邻顶点的平行六面体的体积.

2. 证明：任意三角形的三条中线把三角形分成面积相等的六部分.（提示：用仿射变换把三角形简化为等边三角形的情形.）

3. 证明：任意平行四边形的两条对角线把平行四边形分为面积相等的四部分.

4. (a) 设 P 是平行四边形对角线的交点，证明：过 P 的任意直线把平行四边形分成面积相等的两部分.

   (b) 把这个结果推广到三维情形.

5. 描述三个平面，它们把四面体分为体积相等的六部分.

6. 用平面三角直接证明：由矢量 $\xi=(x_1, x_2)$ 和 $\eta=(y_1, y_2)$ 张成的平行四边形的面积 A 满足等式

$$A^2=\begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix}^2.$$

7. (a) 证明：如果 E_n 中的 m 个矢量 $\alpha_1, \cdots, \alpha_m$ 线性相关，那么由这 m 个矢量张成的平行六面体的 m 维体积等于零.

(b) 叙述并证明这个结果的逆命题.

8. 证明：在正交矩阵群中，所有满足 $|A|=1$ 的矩阵（称为"真"正交矩阵）构成指数为 2 的正规子群.

9. (a) 证明：对应 $A\mapsto|A|$ 把全线性群同态地映射到由非零标量组成的乘法群.

(b) 证明：幺模群是全线性群的正规子群.

(c) 广义幺模群（即所有满足 $|P|=\pm1$ 的矩阵 P）是全线性群的正规子群吗?

10. (a) 证明：如果 A 是以 $\alpha_1, \cdots, \alpha_m$ 为行的任意矩阵，那么 AA^T 是由内积构成的矩阵 $((\alpha_i, \alpha_j))$.

(b) 用(a)证明：如果 $\alpha_1, \cdots, \alpha_m$ 是一组正交矢量，那么

$$|AA^T|=(|\alpha_1|\cdots|\alpha_m|)^2.$$

11. (a) 如果 A 是 $m\times n$ 实矩阵，利用定理 7 的证明来证明 $|AA^T|\geqslant 0$. 证明：当 $m=2$ 时，这个结果就是 § 7.10 定理 18 中的施瓦兹不等式.

(b) 证明：以 $(0,0,0)$，(x_1, y_1, z_1) 和 (x_2, y_2, z_2) 为顶点的三角形面积是 $\frac{1}{2}|AA^T|^{\frac{1}{2}}$，其中 $A=\begin{pmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \end{pmatrix}$.

*(c) 以沿 x 轴，y 轴，z 轴的三个单位线段为棱的四面体的体积为 $\frac{1}{6}$.

证明：以 $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ 为顶点的四面体的体积是 $\frac{1}{6}|BB^T|^{\frac{1}{2}}$，其中 B 是 $3\times n$ 矩阵，它的三行分别为 $\alpha_2-\alpha_1, \alpha_3-\alpha_1, \alpha_4-\alpha_1$.

*(d) 推广到高维的"四面体".

*12. 设 $-K\leqslant a_{ij}\leqslant K\ (i, j=1, \cdots, n)$.

(a) 证明：如果 $\alpha_i=(a_{i1}, \cdots, a_{in})$，那么 $|\alpha_i|\leqslant K\sqrt{n}$.

(b) 证明：$|A|\leqslant|\alpha_1|\cdot|\alpha_2|\cdots|\alpha_n|\leqslant K^n n^{\frac{n}{2}}$(阿达玛(Hadamard)行列式定理)。

§10.4 特征多项式

我们已经看到（§9.2定理5）λ是$n\times n$矩阵A的特征根(特征值)当且仅当矩阵$A-\lambda I$是奇异的．根据定理3，也就是当且仅当$|A-\lambda I|=0$，这就证明了下面的引理．

引理 矩阵A的特征根（特征值）是满足$|A-\lambda I|=0$的标量λ.

在化简矩阵为对角型时，如果化简是可能的，这个引理提供了一个直接的方法．

例 设A是实对称矩阵

$$A=\begin{pmatrix}1 & 3 & 0\\ 3 & -2 & -1\\ 0 & -1 & 1\end{pmatrix},$$

那么按第一行的子式展开$|A-\lambda I|$，

$$|A-\lambda I|=\begin{vmatrix}1-\lambda & 3 & 0\\ 3 & -2-\lambda & -1\\ 0 & -1 & 1-\lambda\end{vmatrix}=-\lambda^3+13\lambda-12.$$

进行因式分解得到$|A-\lambda I|=-(\lambda-1)(\lambda+4)(\lambda-3)$，所以$A$的特征根是1，3，$-4$.（一般地，为了求$3\times3$矩阵的特征根，我们必须解一个三次方程，象§4.4或者§5.5中那样求解.）对每个特征根都存在变换T_A的一个特征矢量.

因为

$$(x, y, z)T_A=(x+3y, 3x-2y-z, -y+z),$$

矢量$\xi=(x,\ y,\ z)$是属于特征根$\lambda=1$的特征矢量当且仅当$x+3y=x$，$3x-2y-z=y$，$-y+z=z$；也就是当且仅当$y=0$和$z=3x$，

于是得到 $\xi=(x, 0, 3x)$. 类似地, $\xi=(x, y, z)$是属于特征根 $\lambda=3$ 的特征矢量当且仅当 $x+3y=3x, 3x-2y-z=3y, -y+z=3z$, 于是 ξ 只能是矢量$(3, 2, -1)$的标量倍. 同样, 属于特征根 $\lambda=-4$的特征矢量是矢量$(-3, 5, 1)$的标量倍. 这三个特征矢量

$$(1, 0, 3), \quad (3, 2, -1), \quad (-3, 5, 1)$$

相互正交, 因此它们线性无关. 以这三个矢量为行的矩阵 P 是非奇异的. 对于由这三个矢量构成的一组新基底, 变换 T_A 是非奇异的, 它对应的矩阵是 PAP^{-1} (参看 § 9.2 定理 3′ 和定理 4). 我们还可以把这组基底标准化, 得到特征矢量的标准正交基

$$\alpha_1=\frac{1}{\sqrt{10}}(1, 0, 3), \ \alpha_2=\frac{1}{\sqrt{14}}(3, 2, -1),$$

$$\alpha_3=\frac{1}{\sqrt{35}}(-3, 5, 1),$$

以这三个矢量为行的矩阵 Q 是正交的, 并且 $QAQ^{-1}=QAQ'$ 是对角矩阵, 其对角线元素为 $1, 3, -4$.

上述 3×3 对称矩阵 A 是二次型 $x^2+6xy-2y^2-2yz+z^2$ 的矩阵. 前面的分析表明, 这个二次型对于标准正交基("主轴")$\alpha_1, \alpha_2, \alpha_3$, 表为对角型 $x^2+3y^2-4z^2$.

一般地, 设 A 是任意 $n\times n$ 矩阵. 因为行列式是一个多项式, 对每一行的元素来说是线性的, 所以行列式 $|A-\lambda I|$ 是未定元 λ 的 n 次多项式:

$$|A-\lambda I|=(-1)^n\lambda^n+b_{n-1}\lambda^{n-1}+\cdots+b_1\lambda+b_0. \qquad (14)$$

我们将定义 A 的特征多项式是 $c_A(\lambda)=|A-\lambda I|$, A 的特征方程是方程 $|A-\lambda I|=0$. 现在将上面的引理重述如下:

定理 10 矩阵 A 的特征根(特征值)是 A 的特征方程的根.

因为任意复多项式至少有一个根, 所以我们得出下面推论:

推论 在复数域上, 一个线性变换至少有一个 (非零) 特征

矢量.

定理 11 相似矩阵有相同的特征多项式.

证明 设两个相似的矩阵是 A 和 $B=P^{-1}AP$. 因为 $|P^{-1}|=|P|^{-1}$, 并且 $|P|$ 是标量, 所以可以交换, 于是由行列式乘法法则得出

$$|P^{-1}AP-\lambda I|=|P^{-1}AP-\lambda P^{-1}IP|=|P^{-1}(A-\lambda I)P|$$
$$=|P^{-1}|\cdot|A-\lambda I|\cdot|P|=|A-\lambda I|.$$

因此我们得到一个推论: $|A-\lambda I|$ 的逐个系数

$$b_0=|A|,$$
$$b_1=\cdots$$
$$\cdots\cdots\cdots$$
$$b_{n-2}=(-1)^n\sum_{i<j}(a_{ii}a_{jj}-a_{ij}a_{ji}),$$
$$b_{n-1}=(-1)^n(a_{11}+\cdots+a_{nn})$$

是矩阵 A 在相似群 $A\longmapsto P^{-1}AP$ 之下的不变量. 关于 b_i 的某些适当的多项式给出另外一些有用的不变量. 在这样的不变量中, 有一个是

$$\sum_{i,j=1}^{n}a_{ij}a_{ji}=\sum_{i=1}^{n}a_{ii}^2+2\sum_{i<j}a_{ij}a_{ji}$$
$$=b_{n-1}^2+(-1)^{n-1}2b_{n-2}.$$

对于对称矩阵的情形, 这个不变量就是 $\sum a_{ij}^2$.

因为 $|A^{\mathrm{T}}-\lambda I|=|(A-\lambda I)^{\mathrm{T}}|=|A-\lambda I|$ (根据定理 1), 我们还有

推论 矩阵 A 和它的转置矩阵 A^{T} 具有相同的特征多项式, 因此具有相同的特征根.

定理 12 对角线元素为 $d_1, d_2, \cdots, d_n$ 的三角形矩阵 T 的特征多项式是

$$|T-\lambda I|=(d_1-\lambda)(d_2-\lambda)\cdots(d_n-\lambda).$$

因为$T-\lambda I$本身也是三角形矩阵，所以根据§10.1的引理2便得到定理的证明. 由此得到推论：对角线元素(可以重复出现)集合是由特征多项式的全体根(有的是重根)组成. 因此对于两个相似的对角矩阵，对角线元素集合和每个对角线元素出现的次数都是一样的. 这可以叙述如下：

推论 两个对角矩阵是相似的当且仅当只是它们的对角线元素的次序不同.

相似性的这个性质给实二次型的正交变换(§9.10)以新的解释. 如果具有矩阵A的二次型XAX^{τ}通过正交变换$Z=XP$化成对角型$\lambda_1 z_1^2+\cdots+\lambda_n z_n^2$，那么这个新二次型的对角矩阵是$D=PAP^{\tau}$. 因为$P$是正交矩阵，所以$P^{\tau}=P^{-1}$，$D=PAP^{-1}$，因此新矩阵$D$与原矩阵$A$是相似的. 所以$D$的特征值$\lambda_1, \cdots, \lambda_n$与已知矩阵$A$的特征值相同. 这就给出下面§9.10定理21的更强的形式：

定理13 任意实二次型XAX^{τ}可以用正交变换化成对角型$\lambda_1 z_1^2+\cdots+\lambda_n z_n^2$，其中系数$\lambda_i$是$A$的特征方程$|A-\lambda I|=(\lambda_1-\lambda)\cdots(\lambda_n-\lambda)=0$的根.

特征方程以及它的根是由矩阵A唯一确定的. 这就证明了对角型本质上的唯一性，并且给出直接计算系数的方法. 知道了系数以后，我们还可以按照上面指出的方法计算出与此相联系的特征矢量作为主轴.

因为我们知道，任意实对称矩阵正交等价于实对角矩阵，所以我们得到

推论 实对称矩阵的所有特征值都是实的.

附注 如果A是对称矩阵，那么属于不同特征值$\lambda_1 \neq \lambda_2$的特征矢量X_1和X_2必然是正交的，因为双线性表达式$X_1 A X_2^{\tau}$可以按两种方法来计算：

$$(X_1A)X_2^{\tau}=\lambda_1(X_1X_2^{\tau}),$$

$$X_1(AX_2^{\tau})=X_1(X_2A)^{\tau}=\lambda_2(X_1X_2^{\tau}).$$

因为 $\lambda_1\neq\lambda_2$, 所以 $X_1X_2^{\tau}$ 一定是零，因此 X_1 与 X_2 正交.

因此，如果 $n\times n$ 对称矩阵 A 有 n 个不同的特征值 $\lambda_1,\cdots,\lambda_n$, 那么任意 n 个相应的特征矢量 $X_1,\cdots,X_n$是正交的，并且以单位矢量$\frac{X_1}{|X_1|},\cdots,\frac{X_n}{|X_n|}$为行的正交矩阵$P$使得 $PAP^{\tau}=PAP^{-1}$ 成为对角矩阵.

习　　题

1. 设D是对角矩阵，其对角线元素为 3, 1, −1, 而P是以(1, 2, −3), (0, −1, 4), (0, 0, 1)为行的三角形矩阵. 计算 $P^{-1}DP$ 的特征方程，并与D的特征方程进行比较.

2. 计算下列各矩阵的特征值和特征矢量

(a) $\begin{pmatrix}-1 & 2 & 2\\ 2 & 2 & 2\\ -3 & -6 & -6\end{pmatrix}$,　(b) $\begin{pmatrix}3 & 2 & 2\\ 1 & 4 & 1\\ -2 & -4 & -1\end{pmatrix}$,

(c) $\begin{pmatrix}4 & 9 & 0\\ 0 & -2 & 8\\ 0 & 0 & 7\end{pmatrix}$.

3. 求二次型 $xy+yz+zx+x+y+z=1$ 的主轴长度.

4. 写出在正交变换之下等价于下面给出的表达式的对角二次型:

(a) $2x^2-11y^2-5z^2+4xy+16yz+20xz$. (提示: 证明所有整数特征值是 9 的倍数.)

(b) $3x^2-y^2-3z^2-t^2-4xz-10yt$.

5. 写出把习题 4 的每个二次型化为等价的对角型的正交变换.

6. 求出 2×2 矩阵的特征值相等的充分必要条件.

7. 求出其特征值为 1 和 −1 的全部 2×2 矩阵.

8. 证明：如果A和B都是方阵，那么矩阵$\begin{pmatrix}A & O\\ C & B\end{pmatrix}$的特征多项式是$A$和$B$的特征多项式的乘积.

9. 证明：(14) 式中的 $b_{n-1}=\pm(a_{11}+\cdots+a_{nn})$.（不变量 $a_{11}+\cdots+a_{nn}$ 称为 A 的迹.）

10. 证明：(14) 式中的 $b_{n-2}=(-1)^n\sum_{i<j}(a_{ii}a_{jj}-a_{ij}a_{ji})$.

11. 证明：对于对称矩阵有 $\Sigma a_{ij}^2=b_{n-1}^2+(-1)^{n-1}2b_{n-2}$.

12. 从定义直接证明：实对称矩阵 A 的所有特征值都是实的.（提示：对于特征矢量 X，证明 $XA(X^*)^T=\lambda X(X^*)^T=\lambda^*X(X^T)^*$，这里 X^* 表示 X 的复共轭.）

13. (a) 证明：埃尔米特矩阵的所有特征值都是实的.

(b) 证明：全体特征矢量张成所有矢量组成的空间.

*14. 证明：每个酉矩阵 U 有一个满足 $\xi U=d\xi$ 的特征矢量 ξ，其中 $|d|=1$.

15. (a) 证明：如果矩阵 A 对于特征值 λ_j 有 r 个线性无关的特征矢量，那么特征多项式 $c_A(\lambda)$ 是 $(\lambda-\lambda_j)^r$ 的一个倍式.

(b) 对任意 r，构造一个 $r\times r$ 矩阵 A，使得 $c_A(\lambda)=(\lambda-\lambda_1)^r$，而对于特征值 λ_1，没有两个线性无关的特征矢量.

16. 通过对线性变换 $X\mapsto XA$ 所进行的如下分析，证明实对称矩阵 A 的主轴定理.

(a) 矩阵 A 有一个长度为 1 的特征矢量 α_1.

(b) 如果把 α_1 选作新的标准正交基的第一个矢量，那么已知变换的新矩阵的第一行和第一列除了第一个元素外其他元素都是零.

(c) 用归纳法继续完成上面的论证.

*17. 证明：椭球 $\Sigma a_{ij}x_ix_j\leqslant 1$ 的体积等于 $\frac{4\pi}{3}|A|^{-\frac{1}{2}}$，其中 $A=(a_{ij})$.（提示：变换到主轴上，并用定理 9.）

§10.5 极小多项式

矩阵在相似变换下标准型的构造依赖于对矩阵或相应的变换所满足的多项式方程的研究．尤其是，设 V 是域 F 上的 n 维矢量空间，$T:V\to V$ 是 V 的线性变换，T 的各次幂 T^m 也是 V 的线性变换．因为线性变换还可以相加或乘上一个标量，所以对每个形为

$f(x)=a_0+a_1x+\cdots+a_kx^k$ (其中系数 $a_i\in F$)的多项式，我们可以考虑相应的关于T的多项式

$$f(T)=a_0I+a_1T+\cdots+a_kT^k. \tag{15}$$

它表示一个线性变换$f(T):V\to V$,特别是常数多项式$f(x)=1$产生恒等变换$I:V\to V$. 因为T的幂是可交换的($T^mT^q=T^qT^m=T^{m+q}$)，所以多项式$f(T)$可以象多项式$f(x)$那样相加和相乘.

类似地,元素在F中的每个$n\times n$矩阵A产生出A的多项式

$$f(A)=a_0I+a_1A+\cdots+a_kA^k, \tag{16}$$

它们还是元素在F中的$n\times n$矩阵. 因为恰好存在n^2个线性无关的F上的$n\times n$矩阵,所以n^2+1个矩阵$I,A,\ \cdots,\ A^{n^2}$一定线性相关，并且这种相关关系提供了一个次数至多为n^2的非零多项式$f(x)$，满足$f(A)=O$. 由于$n\times n$矩阵和V_n的线性变换之间存在同构$A\mapsto T_A$，所以对n维矢量空间V的每个线性变换T也存在一个非零多项式$f(x)$使得$f(T)=O$.

定理 14 对域F上有限维矢量空间V的每个线性变换T，使得$f(T)=O$的F上的多项式$f(x)$是唯一的首一多项式$m(x)$的倍式.

证明 考虑F上满足$f(T)=O$的所有多项式$f(x)$组成的集合M. 我们刚刚看到M包含非零多项式，而且M关于加法、减法和用任意多项式$g(x)$相乘的三种运算是封闭的,因此M是环$F[x]$中的一个理想. 因此根据§3.8定理11,M是由满足$m(T)=O$的次数最小的首一多项式$m(x)$的全体倍式组成.

我们称$m(x)$是T的*极小多项式*. 它是具有下述性质的首一多项式:

$$m(T)=O;\text{由 } f(T)=O \text{ 推出 } m(x)\mid f(x), \tag{17}$$

这里记号$m(x)\mid f(x)$表示在多项式环$F[x]$中$m(x)$整除$f(x)$(象第三章那样). 对$n\times n$矩阵A的极小多项式可做类似的描述，它

等同于 F^n 上相应变换 T_A 的极小多项式. 因为相似的矩阵是同一个线性变换的不同表示, 所以我们有

推论 域 F 上的相似矩阵有相同的极小多项式.

作为一个例子, 我们考虑幂零变换(或幂零矩阵), 即对某个 m 满足 $T^m=O$ 的线性变换 T. 因为 T 满足 $T^m=O$, 所以它的极小多项式是 x^h, 其中 h 是某个整数. 事实上 h 是满足 $T^h=O$ 的最小正整数.

特殊情形是假定 $h=n$. 因为 $T^{h-1}=T^{n-1}\neq O$, 所以存在一个矢量 α 满足 $\alpha T^{n-1}\neq 0$. 我们断言, $\alpha, \alpha T, \alpha T^2, \cdots, \alpha T^{n-1}$ 这 n 个矢量是线性无关的. 如果不然, 则存在线性相关关系 $0=a_0\alpha+a_1\alpha T+\cdots+a_{n-1}\alpha T^{n-1}$, 其系数不全为零. 如果 a_j 是第一个不为零的系数, 那么我们就用 T^{n-j-1} 乘这个方程两边得到

$$0=0T^{n-j-1}=a_j\alpha T^jT^{n-j-1}=a_j\alpha T^{n-1},$$

而这里选取的 α 满足 $\alpha T^{n-1}\neq 0$, 因此 $a_j=0$, 矛盾. 当选取 $\alpha, \alpha T, \cdots, \alpha T^{h-1}$ 这些线性无关矢量作为基底时, T 把每个基矢量变到下一个基矢量, 并把最后一个基矢量变为零矢量, 因此 T 用 $n\times n$ 矩阵表示就是

$$\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}.$$

这个矩阵中非零元素都是 1, 它们分布在紧挨主对角线上面的一条对角线上. 这个矩阵显然是幂零矩阵, 它称为多项式 x^n 的"友矩阵".

更一般地, 对每个 n 次首一多项式

$$g(x)=c_0+c_1x+\cdots+c_{n-1}x^{n-1}+x^n,$$

我们可以构造一个以 $g(x)$ 为极小多项式的 $n\times n$ 矩阵．这个矩阵称为 $g(x)$ 的*友矩阵*，例如 $n=4$ 时，

$$C_g=\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -c_0 & -c_1 & -c_2 & -c_3 \end{pmatrix}; \tag{18}$$

对于任意 n，C_g 的紧挨主对角线上面的一条对角线上的元素都是 1，最后一行的元素是 $-c_0, \cdots, -c_{n-1}$，其余元素都是零．

定理 15 对每个首一多项式 $g(x)$，友矩阵 C_g 有极小多项式 $g(x)$ 和特征多项式 $(-1)^n g(\lambda)$．

证明 设 T 是 F^n 的用形如(18)的友矩阵 C_g 所表示的线性变换．因为这个矩阵的各行分别是 F^n 的单位矢量 $\varepsilon_1, \cdots, \varepsilon_n$ 的变换式的坐标，所以我们有

$$\varepsilon_1 T=\varepsilon_2, \cdots, \varepsilon_{n-1} T=\varepsilon_n,$$
$$\varepsilon_n T=-c_0\varepsilon_1-\cdots-c_{n-1}\varepsilon_n.$$

换句话说，矢量 $\varepsilon_1, \varepsilon_1 T, \cdots, \varepsilon_1 T^{n-1}$ 是 F^n 的一组基底，所以任意矢量 ξ 可以唯一地写成

$$\xi=a_0\varepsilon_1+a_1\varepsilon_1 T+\cdots+a_{n-1}\varepsilon_1 T^{n-1}=\varepsilon_1 f(T), \tag{19}$$

这里 $f(x)=a_0+a_1x+\cdots+a_{n-1}x^{n-1}$ 是次数至多为 $n-1$ 的多项式．进一步有 $\varepsilon_1 T^n=-c_0\varepsilon_1-\cdots-c_{n-1}\varepsilon_1 T^{n-1}$，所以 $\varepsilon_1 g(T)=0$．因此对任意矢量 ξ，

$$\xi g(T)=\varepsilon_1 f(T) g(T)=\varepsilon_1 g(T) f(T)=0,$$

这就断言，T 满足首一多项式方程 $g(T)=O$．对于任意次数较低的多项式 $f(x)\neq 0$，由(19)式可知 $\varepsilon_1 f(T)=\xi\neq 0$，因此 $f(T)\neq O$．于是 $g(x)$ 确实是 C_g 的极小多项式．

C_g 的特征多项式是通过把行列式 $|C_g-\lambda I|$ 按最后一行的子式展开求得的．因为 $-c_k$ 的子式是三角形矩阵，它的对角线元素

有 k 个是 $-\lambda$，其他都是 1，所以 $|C_g-\lambda I|$ 确实等于 $(-1)^n g(\lambda)$，这里出现符号 $(-1)^n$ 是因为任意 $n\times n$ 矩阵的特征多项式的首项都是 $(-1)^n\lambda^n$。

习　题

1. (a) 证明：任意满足 $X^2=O$ 的 2×2 矩阵 X 相似于矩阵 $\begin{pmatrix}0&1\\0&0\end{pmatrix}$ 或者 X 是 $\begin{pmatrix}0&0\\0&0\end{pmatrix}$.

(b) 对 3×3 矩阵证明相应的结果.

2. 证明：每个实 2×2 矩阵，如果它的行列式是负的，那么它同对角矩阵相似. 给出几何解释.

3. (a) 对任意 $n\times n$ 非奇异矩阵 P，证明：对应 $A\mapsto PAP^{-1}$ 是所有 $n\times n$ 矩阵 A 的代数的自同构.

(b) 根据(a)直接证明：相似矩阵有相同的极小多项式.

4. 证明：任意对角矩阵的特征多项式是它的极小多项式的倍式. 什么时候两者相同?

*5. (a) 证明：每个行列式为负的 2×2 实正交矩阵表示一个刚体反射. (提示：见习题 2，或 § 9.4.)

(b) 证明：每个行列式为正的 2×2 正交矩阵表示一个刚体旋转.

*6. (a) 证明：任意 3×3 实矩阵 A 有一个实特征矢量.

(b) 证明：任意 3×3 正交矩阵，在基底的正交变换下，它相似于形为 $\begin{pmatrix}\pm1&0\\0&B\end{pmatrix}$ 的矩阵，这里 B 是 2×2 正交矩阵.

(c) 用习题 5 证明：如果 A 是 3×3 正交矩阵，并且 $|A|>0$，那么 A 有一个特征值是 1，而且 A 表示刚体旋转. (这就是欧拉定理.)

*7. 证明：如果 λ 是 A 的特征值，$q(x)$ 是任意多项式，那么 $q(\lambda)$ 是 $q(A)$ 的特征值.

*8. (a) 证明：矩阵

$$C=\begin{pmatrix}0&1&0&0\\0&0&1&0\\0&0&0&1\\1&0&0&0\end{pmatrix}$$

的特征值为±1, ±i, 即复四次单位根.

(b) C 的复特征矢量是什么?

(c) C 与什么样的复对角矩阵相似?

*9. 一个 $n\times n$ 矩阵, 如果它满足对一切 $i, j, a_{ij}=a_{i+1,j+1}$(其中下标 i, j 都取模 n 整数), 则称这个矩阵为轮换矩阵. 证明: 任意轮换矩阵的特征值 $\lambda_1, \cdots, \lambda_n$ 是

$$\lambda_p=a_{11}+a_{12}\omega^p+\cdots+a_{1n}\omega^{(n-1)p},$$

式中 ω 是 n 次本原单位根. (提示: 利用习题7和习题8.)

§10.6 凯莱-哈密顿定理

我们现在将要证明, 每个方阵 A 都满足它的特征方程, 也就是说, A 的极小多项式整除 A 的特征多项式.

用矩阵多项式或者 λ-矩阵的概念很容易证明上述性质. λ-矩阵 $A-\lambda I$, 是指以 λ 多项式为元素的矩阵. 把 λ 的同次幂的项合并, 我们可以把任意非零 λ-矩阵 $B(\lambda)$ 写成下面形式

$$B(\lambda)=B_0+\lambda B_1+\cdots+\lambda^r B_r,$$

这里 B_i 是常数矩阵, $B_r\neq O$. (等式意味着两边的矩阵各个元素中 λ 的每个相应系数都相等.)

引理 如果 $C=B(\lambda)(A-\lambda I)$ 是常数矩阵, 那么 $C=O$.

证明 展开 $B(\lambda)(A-\lambda I)$, 我们得到

$$-\lambda^{r+1}B_r+\sum_{k=1}^{r}\lambda^k(B_kA-B_{k-1})+B_0A,$$

如果它是常数矩阵, 意味着 $B_r=O$, $B_kA-B_{k-1}=O$, $k=1, \cdots, r$. 故得 $B(\lambda)=O$. 现在结论显然成立.

定理 16(凯莱-哈密顿) 每个方阵满足它的特征方程.

这意味着, (14)的特征多项式 $f(\lambda)=|A-\lambda I|$ 中, 如果 λ 的每个幂 λ^i 用矩阵 A 的同次幂 A^i 代替, (并且 λ^0 用 $A^0=I$ 代替), 其结果是零:

$$b_0 I + b_1 A + \cdots + b_{n-1} A^{n-1} + (-1)^n A^n = O. \tag{20}$$

证明 在矩阵 $A-\lambda I$ 中，每个元素都是 λ 的线性多项式，所以它的非零子式也是 λ 的 $n-1$ 次或小于 $n-1$ 次的多项式. $A-\lambda I$ 的伴随矩阵 C 的每个元素是这样的子式，所以这个伴随矩阵可以写成 n 个矩阵的和，这 n 个矩阵分别包含 λ 的固定的幂 $\lambda^0, \lambda^1, \cdots, \lambda^{n-1}$. 换句话说，伴随矩阵 $C=C(\lambda)$ 是 λ-矩阵 $C=C(\lambda)=\sum\lambda^i C_i$. 根据(10)式，$A-\lambda I$ 与它的伴随矩阵的乘积是

$$C(\lambda)(A-\lambda I) = |A-\lambda I| I = f(\lambda) I, \tag{21}$$

这里 $f(\lambda)$ 是特征多项式.

现在注意，由熟悉的因式分解公式

$$A^i - \lambda^i I = (A^{i-1} + \lambda A^{i-2} + \cdots + \lambda^{i-1} I)(A-\lambda I),$$

再利用特征多项式(14)的系数 b_i，我们得到

$$\begin{aligned} f(A) - f(\lambda) I &= \sum_{i=0}^{n} b_i A^i - \sum_{i=0}^{n} b_i \lambda^i I \\ &= \sum_{i=1}^{n} b_i (A^i - \lambda^i I) \\ &= \sum_{i=1}^{n} b_i (A^{i-1} + \lambda A^{i-2} + \cdots \\ &\qquad + \lambda^{i-1} I)(A - \lambda I), \end{aligned}$$

这里 $f(A)$ 是特征多项式 $f(\lambda)$ 中的 λ 换成 A 而得到的. 上式即

$$f(A) - f(\lambda) I = -G(\lambda)(A-\lambda I), \tag{22}$$

其中 $G(\lambda)$ 是一个新的 λ- 矩阵. 如果我们把(22)与(21)相加，便得

$$[C(\lambda) - G(\lambda)](A-\lambda I) = f(A),$$

其中 $f(A)$ 是常数矩阵. 根据引理，这就推出 $f(A)=O$.

习　题

1. 通过直接代入证明：每个 2×2 矩阵满足它的特征方程.

2. 证明：如果 A 是非奇异矩阵，并有特征多项式(14)，那么 A 的伴随矩阵由

$$-[b_1I+b_2A+\cdots+b_{n-1}A^{n-2}+(-1)^nA^{n-1}]$$

给出.

3. 根据习题 2 的表示法证明：A^{-1} 的特征多项式为

$$(-1)^n\left[\lambda^n+\frac{b_1}{|A|}\lambda^{n-1}+\cdots+\frac{(-1)^n}{|A|}\right].$$

4. (a) 通过直接计算证明：关于严格三角形矩阵的凯莱-哈密顿定理.

　*(b) 对三角形矩阵回答同样问题.

5. (a) 用直接计算证明：(18)式的 4×4 友矩阵 C_g 满足它的特征方程.

　(b) 对 n 次多项式的友矩阵做同样证明.

§10.7　不变子空间与可约性

如果一个线性变换 T 满足一个可以因式分解的多项式方程，那么表示 T 的矩阵常常能够相应地简化. 例如，我们假定 T 满足 $T^2=I$(周期是 2)，在基域 F 中 $1+1\neq0$，所以 $(T-I)(T+I)=O$ 的两个因子是互素的. T 的特征矢量包含 $T+I$ 的值域中的全部非零矢量 $\eta=\xi(T+I)$，这因为

$$(\xi(T+I))T=\xi(T^2+T)=\xi(T+I).$$

这些矢量 η 是属于特征值 1 的. 同样地，$T-I$ 的值域中的所有非零矢量是属于特征值 -1 的特征矢量，这因为

$$(\xi(T-I))T=\xi(T^2-T)=\xi(I-T)=-\xi(T-I).$$

而因为 $1+1\neq0$，所以任意矢量 ξ 可以写成两个矢量的和

$$\xi=\frac{1}{2}[\xi(T+I)-\xi(T-I)].$$

因此属于特征值 ±1 的特征矢量张成这个空间. 于是根据 §9.2 定

理 4, T 可以用对角线元素是 ± 1 的对角矩阵表示.

特别地,如果对角线元素都是 1, 则 T 是恒等变换, 并且 T 的极小多项式是 $x-1$; 如果对角线元素都是 -1, 则 T 的极小多项式是 $x+1$; 如果对角线元素既有 1 又有 -1, 则 T 的极小多项式是 x^2-1. 这个分析是下面定理的特殊情形.

定理 17 *如果线性变换 $T: V \to V$ 的极小多项式 $m(x)$ 在 V 的基域 F 上可以分解因式为 $m(x)=f(x)g(x)$, 其中 $f(x)$ 和 $g(x)$ 是首一多项式, 并且互素, 那么 V 中任意矢量可以唯一地表示成和*

$$\xi=\eta+\zeta, \quad \eta f(T)=0, \quad \zeta g(T)=0. \tag{23}$$

证明 因为 f 和 g 是互素的, 所以由欧几里得算法给出系数在 F 中的两个多项式 $h(x)$ 和 $k(x)$, 使得

$$1=h(x)f(x)+k(x)g(x). \tag{24}$$

将 T 代入 x, 得到 $I=h(T)f(T)+k(T)g(T)$. 于是, 对任意矢量 ξ, 有

$$\xi=\xi I=\eta+\zeta, \quad \eta=\xi k(T)g(T), \quad \zeta=\xi h(T)f(T).$$

因为 $\eta f(T)=\xi k(T)g(T)f(T)=\xi k(T)m(T)=0$, 类似地, $\zeta g(T)=0$, 所以这就是所要求的分解.

分解式(23)是唯一的. 因为如果 $\xi=\eta_1+\zeta_1=\eta_2+\zeta_2$ 是两个分解式, 那么 $\alpha=\eta_1-\eta_2=\zeta_2-\zeta_1$ 是满足 $\alpha f(T)=0$ 和 $\alpha g(T)=0$ 的矢量. 因此由(24), 有

$$\alpha I=\alpha h(T)f(T)+\alpha k(T)g(T)=0,$$

于是 $\eta_1=\eta_2$, $\zeta_1=\zeta_2$.

定理 17 也可按另一种方式叙述. 子空间 S_1 是由所有满足 $\eta f(T)=0$ 的矢量 η 组成, S_2 是由所有满足 $\zeta g(T)=0$ 的矢量 ζ 组成. 这就是说, S_1 是 $f(T)$ 的零空间, S_2 是 $g(T)$ 的零空间. 此外, 按照 § 7.8 中的定义, V 是子空间 S_1 与 S_2 的直和. 这两个子空间, 每一个都被 T 映射到自身. 这样, 按下述一般定义, 它是一个

“不变”子空间.

矢量空间V的子空间S在线性变换$T:V\to V$之下，如果由$\xi\in S$推出$\xi T\in S$，则称S为T之下的不变子空间. 这时，对应$\xi\longmapsto\xi T$称为T在S上的导出变换.

显然，如果S在T之下是不变子空间，而$h(x)$是任意多项式，那么S在$h(T)$之下也是不变子空间.

在定理 17 中，对每个$\eta\in S_1$，有$\eta f(T)=0$；因此，如果T_1是T在S_1上的导出线性变换，那么T_1的极小多项式是$f(x)$的因子$f_1(x)$. 类似地，S_2上的导出变换T_2的极小多项式是$g(x)$的因子$g_2(x)$. 所以对任意表示成(23)的矢量ξ，有

$$\begin{aligned}\xi f_1(T)g_2(T)&=[\eta f_1(T)]g_2(T)+[\zeta g_2(T)]f_1(T)\\&=0+0=0,\end{aligned}\tag{25}$$

因此乘积$f_1(x)g_2(x)$可被极小多项式$m(x)=f(x)g(x)$整除. 因为f和g是互素的，这就证明了，$f(x)$可整除$f_1(x)$，$g(x)$可整除$g_2(x)$. 但是$f_1(x)$也可整除$f(x)$，所以$f_1=f$，同样有$g_2=g$. 于是我们就得到下面的结果.

定理 17′ 如果S_1和S_2分别是定理 17 中$f(T)$和$g(T)$的零空间，那么V是S_1与S_2的直和，并且T在S_1与S_2上的导出变换T_1和T_2分别有极小多项式$f(x)$和$g(x)$.

可以通过很多方式产生不变子空间. 比如，设$f(x)$是任意多项式，那么变换$f(T):V\to V$的值域——即所有矢量$\xi f(T)$(其中$\xi\in V$)组成的集合——是在T之下的不变子空间，这是因为$\xi f(T)T=(\xi T)f(T)$也是在这个值域中. 一类特殊的不变子空间是由一个矢量生成的循环子空间. 我们现在给出定义.

已知变换$T:V\to V$和V中一个矢量α，显然，V的任意子空间，如果它包含α，并且在T之下是不变子空间，那么它一定包含α的由T的多项式$f(T)$作用而得的所有变换式$\alpha f(T)$. 但是，所有

这种变换式的集合 Z_α 是一个包含 α 的不变子空间，我们称它是由 α 生成的 T-循环子空间.

我们现在考虑 α 在 T 的逐次幂之下的一列变换式 $\alpha=\alpha I, \alpha T, \alpha T^2, \cdots$. 显然，存在第一个与它前面的变换式线性相关的变换式 αT^d. 那么我们有

$$\alpha T^d+c_{d-1}\alpha T^{d-1}+\cdots+c_0\alpha I=\alpha m_\alpha(T)=0, \tag{26}$$

这里 $\alpha, \alpha T, \cdots, \alpha T^{d-1}$ 是线性无关的. 于是 $m_\alpha(x)=x^d+c_{d-1}x^{d-1}+\cdots+c_0$ 是变换 T_α 的极小多项式，其中 T_α 是 T 在 T-循环子空间 Z_α 上的导出变换；多项式 $m_\alpha(x)$ 称为 α 的 T-阶. 注意，T 把 Z_α 的基矢量 $\alpha, \alpha T, \cdots, \alpha T^{d-1}$ 中的每一个矢量都映射到它的下一个矢量，而 αT^{d-1} 除外，它被映射到

$$\alpha T^d=-c_0\alpha-c_1\alpha T-\cdots-c_{d-1}\alpha T^{d-1}. \tag{27}$$

对于 Z_α 的这组基底 $\alpha, \alpha T, \cdots, \alpha T^{d-1}$，表示 T_α 的矩阵的行是基矢量的变换式的坐标 $(0, 1, 0, \cdots, 0), (0, 0, 1, \cdots, 0), \cdots, (0, \cdots, 0, 1), (-c_0, \cdots, -c_{d-1})$. 这个矩阵恰恰是多项式 $m_\alpha(x)$ 的友矩阵，所以我们就证明了

定理 18 *T 在 T-循环子空间 Z_α 上导出的具有 T-阶 $m_\alpha(x)$ 的导出变换可以用 $m_\alpha(x)$ 的友矩阵来表示.*

反过来，n 次首一多项式 f 的友矩阵 C_f 表示变换 $T=T_{C_f}: F^n\to F^n$，它把 F^n 的每个单位矢量 ε_i 变换到下一个单位矢量 ε_{i+1}，把最后一个单位矢量 ε_n 变换到 $\varepsilon_1 T^n$. 因此，象在(19)式那样，整个空间 F^n 是由 ε_1 生成的 T-循环子空间，并具有 T-阶 $f(x)$.

定理 19 *如果变换 $T: V\to V$ 的极小多项式是 $m(x)$，那么 V 中每个矢量 α 的 T-阶是 $m(x)$ 的一个因子.*

证明 因为 $m(T)=O, \alpha m(T)=0$，所以，由(26)式，$m(x)$ 是 α 的 T-阶 $m_\alpha(x)$ 的倍式.

推论 *V 中两个矢量 α 和 β 张成相同的 T-循环子空间 $Z_\alpha=$*

Z_β 当且仅当 $\beta=\alpha g(T)$，这里多项式 $g(x)$ 与 α 的 T-阶 $m_\alpha(x)$ 是互素的.

证明留作习题(习题 8)。

习　题

1. (a) 证明：满足 $A^2=-I$ 的任意 2×2 实矩阵 A 与矩阵 $\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}$ 相似.

(b) 证明：不存在 3×3 实矩阵满足 $A^2=-I$.

(c) 对于满足 $A^2=-I$ 的 4×4 实矩阵 A，有什么结果?

2. (a) 证明：满足 $T^2=T$ 的任意"幂等"线性变换 T 的值域和零空间是补子空间(参看 §7.8 的(35)式).

(b) 证明：任意两个秩相同的幂等矩阵相似.(提示：用(a)的结果.)

3. (a) 对所有满足 $A^3=I$ 的 3×3 复矩阵进行分类.

(b) 对 3×3 实矩阵解答同样的问题.

4. 每一个平面切变换满足 $A^2+I=A+A$. 求满足这个方程的 2×2 矩阵的标准型.(提示：考虑形式 $A-I$.)

*5. 当标量域具备什么条件时，具有极小多项式 x^2+x-2 的矩阵与 2×2 对角矩阵相似.

6. (a) 在定理 17 中，证明：$g(T)$ 的值域同 $f(T)$ 的零空间相等.

(b) 证明：如果 $f(T)g(T)=O$，其中 $f(x)$ 和 $g(x)$ 是互素的多项式，那么，即使 $f(x)g(x)$ 不是 T 的极小多项式，定理 17 的结论也成立.

7. 证明：矢量 α 的 T-阶是满足 $\alpha f(T)=0$ 的次数最低的首一多项式 $f(x)$.

8. 证明定理 19 的推论.

9. 证明：已知变换 $T:V\to V$，V 中的矢量 α 与 β 具有互素的 T-阶 $f(x)$ 与 $g(x)$，那么 $\alpha+\beta$ 具有 T-阶 $f(x)g(x)$.

*10. 证明：T-循环空间的每个不变子空间本身是 T-循环空间.(提示：考虑循环群相应的性质.)

11. 如果 $f(T)=O$，而 $f(x)$ 和 $g(x)$ 是互素的，那么 T 和 $g(T)$ 具有相同的循环子空间.

§10.8 第一分解定理

在证明定理 17 和定理 17′ 时用过的结构可以用来分解一般线性变换为"准素"分支，这些分支的极小多项式是不可约多项式的幂．在这个分解中，k 个子空间的直和的概念起着重要的作用．

定义 我们称矢量空间 V 是它的子空间 $S_1, \cdots, S_k$ 的直和（用符号表示就是 $V = S_1 \oplus \cdots \oplus S_k$），是指 V 中每个矢量 ξ 可以唯一地表示为

$$\xi = \eta_1 + \cdots + \eta_k \quad (\eta_i \in S_i,\ i = 1, \cdots, k). \tag{28}$$

同 §7.8 的定理 16 完全一样，我们可以证明

定理 20 如果 V 有子空间 $S_1, \cdots, S_k$，其中每个 S_i 的维数是 n_i，并有基底 $\alpha_{i1}, \cdots, \alpha_{in_i}$，那么 V 是 $S_1, \cdots, S_k$ 的直和当且仅当

$$\alpha_{11}, \cdots, \alpha_{1n_1};\ \alpha_{21}, \cdots, \alpha_{2n_2};\ \cdots;\ \alpha_{k1}, \cdots, \alpha_{kn_k} \tag{29}$$

是 V 的一组基底．

由此得出，V 的维数是直和被加项 S_i 的维数之和 $n_1 + \cdots + n_k$．

推论 如果 V 是由子空间 $S_1, \cdots, S_k$ 张成，并且

$$d[V] = d[S_1] + \cdots + d[S_k],$$

那么 V 是 $S_1, \cdots, S_k$ 的直和．

如果空间 V 可以表示成变换 T 之下的真不变子空间的直和，那么我们称线性变换 $T: V \to V$（或对应于 T 的矩阵）是完全可约的．

定理 21 如果 V 是不变子空间 $S_1, \cdots, S_k$ 的直和，已知变换在每个子空间上的导出变换用矩阵 B_i 表示，那么 T 在 V 上可以用矩阵

$$B=\begin{pmatrix} B_1 & O & \cdots & O \\ O & B_2 & \cdots & O \\ \cdots & \cdots & \cdots & \cdots \\ O & O & \cdots & B_k \end{pmatrix} \tag{30}$$

表示.

这个矩阵B沿对角线排列着矩阵块$B_1, \cdots, B_k$, 其他位置的元素都是零,称B为矩阵$B_1, \cdots, B_k$的*直和*. 可以看出, B的任意多项式$f(B)$是$f(B_1), \cdots, f(B_k)$的直和.

证明 对每个不变子空间S_i,选取一组基底$\alpha_{i1}, \cdots, \alpha_{in_i}$,所以$B_i$是变换$T$在$S_i$上对于这组基底的矩阵表示, 那么把这些基矢量合并在一起构成整个空间的一组基底 (29). 而且T把基矢量$\alpha_{i1}, \cdots, \alpha_{in_i}$变到第$i$个子空间的矢量, 因此变换$T$对于基底(29)可用上述直和矩阵(30)来表示. 证毕

现在考虑在基域上把T的极小多项式$m(x)$分解成不同的首一不可约多项式$p_i(x)$的幂的乘积,写成形式

$$m(x)=p_1(x)^{e_1}\cdots p_k(x)^{e_k}, \qquad e_i>0 \tag{31}$$

因为不同的$p_i(x)^{e_i}$是互素的,所以反复运用定理17′就得到

定理22 *如果线性变换$T:V\to V$的极小多项式在基域F上可以分解成首一不可约因子$p_i(x)$的乘积((31)式),那么V是不变子空间$S_1, \cdots, S_k$的直和, 其中S_i是$p_i(T)^{e_i}$的零空间. T在S_i上的导出变换T_i有极小多项式$p_i(x)^{e_i}$.*

这就是我们的第一分解定理. 这些子空间S_i称为V的在T之下的"准素分支". 它们由T唯一确定, 因为分解式 (31) 是唯一的.

一个重要的特殊情形是

推论 *元素在F中的矩阵A在F上相似于对角矩阵当且仅当A的极小多项式$m(x)$是F上不同线性因子的乘积.*

证明 设 $T=T_A:F^n\to F^n$ 是对应于 A 的变换. 如果

$m(x)=(x-\lambda_1)\cdots(x-\lambda_k)$, $\lambda_1,\cdots,\lambda_k$ 是不同的标量, (32)

这个定理表明 V 是空间 $S_1,\cdots,S_k$ 的直和,其中 S_i 是由满足 $\eta_i T=\lambda_i\eta_i$ 的所有矢量 η_i 组成,也就是说,由属于特征值 λ_i 的所有特征矢量组成. S_i 的任意基底必由这样的特征矢量组成,所以变换 T 在 S_i 上的矩阵表示是 $\lambda_i I$. 象(29)式那样把这些基底合并在一起,我们可以用一个对角矩阵来表示 T,这个对角矩阵的对角线上的元素是 $\lambda_1,\cdots,\lambda_k$.

反过来,如果 D 是任意对角矩阵,它的不同的对角线元素是 $c_1,\cdots,c_k$,那么用乘积 $f(D)=(D-c_1I)\cdots(D-c_kI)$ 表示的变换把每个基矢量映射到 0,因此 $f(D)=O$. D 的极小多项式或者与 D 相似的任意其他矩阵的极小多项式是乘积 $(x-c_1)\cdots(x-c_k)$ 的因子,因此是不同线性因子的乘积.

习 题

1. 证明定理 20.

2. 在定理 22 中,设 $q_i(x)=\dfrac{m(x)}{p_i(x)^{e_i}}$, 证明:那里的子空间 S_i 是 $q_i(T)$ 的值域.

*3. 不用定理 17′,直接证明定理 22.

4. 证明:如果 $n\times n$ 矩阵 A 相似于对角矩阵 D,那么 D 的对角线元素 λ_i 出现的次数等于属于特征值 λ_i 的特征矢量的集合的维数.

5. 证明:两个矩阵 B_1 与 B_2 的直和的极小多项式是 B_1 与 B_2 的极小多项式的最小公倍式.

6. 证明:矩阵 A 的极小多项式能够分解成线性因子当且仅当 A 的特征多项式可以同样地分解.

*7. 设 A 是复矩阵,它的极小多项式 $m(x)=(x-\lambda_1)^{e_1}\cdots(x-\lambda_r)^{e_r}$ 等于它的特征多项式. 证明:矩阵 A 同 r 个 $e_i\times e_i$ 三角形矩阵 B_i 的直和相似,B_i 的形状如:

$$B_i=\begin{pmatrix} \lambda_i & 1 & & & 0 \\ & \lambda_i & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda_i & 1 \\ 0 & & & & \lambda_i \end{pmatrix}.$$

*8. 证明：如果 $m(x)$ 是 T 的极小多项式，那么存在一个矢量 α，它的 T-阶恰好是 $m(x)$.（提示：利用 § 10.7 的习题 9，首先考虑 $m(x)=p(x)^e$ 的情形，其中 $p(x)$ 是不可约多项式.）

§ 10.9 第二分解定理

下面我们将指出，线性变换 $T:V\to V$ 的准素分支 S_i 本身是 T-循环子空间的直和. 在证明这个命题时，我们将用到矢量空间 V 对于子空间 S 的商空间 V/S 的概念. 我们回忆一下（§ 7.12），商空间 $V'=V/S$ 的元素是 S 的陪集 $\xi+S$，并且由 $\xi P=\xi+S$ 给出的射影 $P:V\to V/S=V'$ 是一个线性变换. 特别是，对给定的 $T:V\to V$，如果子空间 S 是 T 之下的不变子空间，那么在公式

$$(\xi+S)T'=\xi T+S \tag{33}$$

中，$\xi T+S$ 不依赖于 $\xi'=\xi+S$ 中的代表元 ξ 的选取，这是因为如果选取另外的代表元 $\eta=\xi+\zeta$，那么由于 $\zeta\in S$，推出 $\zeta T\in S$，所以有

$$\eta T+S=\xi T+\zeta T+S=\xi T+S.$$

因此由(33)式定义的变换 $T':V'\to V'$ 是单值的；容易验证 T' 也是线性的. 我们称 T' 是 $V/S=V'$ 上的由 T 导出的变换. 而且对于 T 的任意多项式 $f(T)$，利用 § 7.12 的公式，由(33)得到

$$(\xi+S)f(T')=\xi f(T)+S. \tag{34}$$

特别是，由 $f(T)=O$ 推出 $f(T')=O'$ 在 V' 中，所以 V' 中 ξ' 的 T'-阶整除 V 中 ξ 的 T-阶.

我们现在准备证明第二分解定理.

定理 23 如果线性变换 $T:V\to V$ 有极小多项式 $m(x)=p(x)^e$，

它是V的标量域F上首一不可约多项式$p(x)$的幂，那么V是T-循环子空间Z_i的直和

$$V=Z_1\oplus\cdots\oplus Z_r, \tag{35}$$

其中子空间$Z_1, Z_2, \cdots, Z_r$分别具有T-阶

$$p(x)^{e_1},\ p(x)^{e_2},\ \cdots,\ p(x)^{e_r},\ e=e_1\geqslant e_2\geqslant\cdots\geqslant e_r. \tag{36}$$

V的任意T-循环子空间的直和表达式有相同个数的分支子空间和相同的T-阶集合(36).

证明 直和分解的存在性可以通过对V的维数n用归纳法来证明. 当$n=1$时，V本身就是循环子空间，于是直接可得结论.

当$n>1$时，我们有$p(T)^e=O$，而$p(T)^{e-1}\neq O$，因此V包含一个满足$\alpha_1 p(T)^{e-1}\neq 0$的矢量$\alpha_1$. 所以$\alpha_1$的$T$-阶是$p(x)^e$，并且$\alpha_1$生成$T$-循环子空间$Z_1$. 因为$Z_1$是$T$之下的不变子空间，所以$T$导出一个在$V'=V/Z_1$上的线性变换$T'$. 因为显然有$p(T')^e=O'$，所以在$V'$上$T'$的极小多项式是$p(x)^e$的因子，我们可以对$d[V/Z_1]=d[V]-d[Z_1]$用归纳法，把$V/Z_1$分解成$T'$-循环子空间$Z_2', \cdots, Z_r'$的直和，这些子空间的$T'$-阶是

$$p(x)^{e_2'},\quad \cdots,\quad p(x)^{e_r'},\quad e\geqslant e_2'\geqslant\cdots\geqslant e_r'.$$

引理1 如果α_i'生成T'-循环子空间Z_i' $(i=2, \cdots, r)$，那么陪集α_i'包含代表元α_i，而α_i的T-阶是α_i'的T'-阶.

证明依赖于下述事实：α_1的T-阶$p(x)^e$是V的每个元素的T-阶的倍式. 特别设$p(x)^d$是α_i'的T'-阶，所以对α_i'的任意代表元$\eta=\eta_i$，$\eta p(T)^d=\alpha_1 f(T)$在$\alpha_1$生成的$T$-循环子空间中. 那么

$$0=\eta p(T)^e=\alpha_1 f(T)p(T)^{e-d}.$$

因为α_1有T-阶$p(x)^e$，所以上式推出$p(x)^e | f(x)p(x)^{e-d}$，因此$f(x)=g(x)p(x)^d$，其中$g(x)$是某一多项式. 我们现在将证明$\alpha_i=\eta-\alpha_1 g(T)$的$T$-阶$p(x)^d$同$\alpha_i'$的$T'$-阶相等. 这是引理所要求的. 因为$\alpha_i$的$T$-阶是$\alpha_i'=\alpha_i+Z_1$的$T'$-阶$p(x)^d$的倍式，所以

只须注意

$$[\eta-\alpha_1 g(T)]p(T)^d=\eta p(T)^d-\alpha_1 f(T)=0.$$

证完引理 1, 我们设 Z_i 是由 α_i 生成的 T-循环子空间. 那么 $d[Z_i]=d[Z_i']$,这因为这两个维数都等于 α_i' 的公共 T-阶 $p(x)^{e_i'}$ 的次数. 因此

$$d[V]-d[Z_1]=d[V/Z_1]=d[Z_2]+\cdots+d[Z_r]. \qquad (37)$$

通过选取基底可以得到,子空间 $Z_1, \cdots, Z_r$ 张成 V; 因此根据(37)和定理 20 的推论得到, V 是直和 $V=Z_1\oplus\cdots\oplus Z_r$, 正如断言所述.

剩下需要证明出现在分解式(36)中指数的唯一性, 这只须证明这些指数是由 T 和 V 确定的. 通过对这些子空间维数的计算, 就可做到这一点. 例如, 如果 d 表示 $p(x)$的次数, 那么循环子空间 Z_i 的维数是 de_i, 因此整个空间 V 的维数是 $d(e_1+\cdots+e_r)$. 还可以看出, 对任意整数 s, Z_i 在 $p(T)^s$ 之下的象 $Z_ip(T)^s$ 是由 $\beta_i=\alpha_ip(T)^s$ 生成的循环子空间. 当 $e_i>s$ 时, 它的维数是 $d(e_i-s)$, 当 $e\leqslant s$ 时, 它的维数是零.

V 的任意矢量 ξ 有唯一的表达式

$$\xi=\eta_1+\cdots+\eta_r \quad (\eta_i\in Z_i,\ i=1, \cdots, r).$$

因此在 $p(T)^s$ 的值域 $Vp(T)^s$ 中, 任意矢量有唯一表达式

$$\xi p(T)^s=\eta_1 p(T)^s+\cdots+\eta_r p(T)^s, \qquad (38)$$

其中分量 $\eta_ip(T)^s$ 在空间 $Z_ip(T)^s$ 中. 整数 s 确定一个整数 t, 使得

$$e_1>s, \cdots, e_t>s, \quad e_{t+1}\leqslant s$$

(或者, 当$e_r>s$ 时, $t=r$). 因此由(38)式, $Vp(T)^s$是由$\beta_i=\alpha_ip(T)$ 生成的循环子空间 $Z_{\beta_i}(i=1, \cdots, t)$的直和, 并且它的维数是

$$d[Vp(T)^s]=d[(e_1-s)+\cdots+(e_t-s)]. \qquad (39)$$

等式左边的维数由 V 和 T 确定, 它们又依次确定 e_i 如下. 首先取 $s=e-1=e_1-1$, 则由(39)式确定出等于 e 的 e_i 的个数; 其次取

$s=e-2$, 则由(39)式确定出等于 $e-1$ 的 e_i(如果有的话)的个数, 等等. 这就证明了指数 $e_1, \cdots, e_r$ 的不变性, 从而完成了定理 23 的证明.

习　题

1. 证明: 如果矢量空间 V 是由矢量 $\alpha_1, \cdots, \alpha_n$ 生成的 T-循环子空间张成的, 那么 T 的极小多项式是 $\alpha_1, \cdots, \alpha_n$ 的 T-阶的最小公倍式.

2. 求 § 8.5 习题 3 中的矩阵 B 的极小多项式.

3. 详细证明: 如果变换 $T: V\to V$ 是线性的, 子空间 Z 是 T 之下的不变子空间, 那么 $T': V/Z\to V/Z$ 是线性的.

4. 根据(37)证明: $Z_1, \cdots, Z_r$ 张成空间 V.

§ 10.10　有理标准型与若当标准型

用定理 20 和定理 23 容易得到矩阵在相似变换之下的标准型. 我们只须对于循环子空间上的变换给出标准型.

定理 21 提供了这样的标准型. 如果 A 是任意 $n\times n$ 矩阵, 那么在每个循环子空间中适当选取一组基底, 在这个子空间上, T_A 用友矩阵表示. 把所有这些基底合并起来产生 F^n 的一组基底, 对于这组基底, T_A 可用这些友矩阵的直和来表示. 定理 20 和定理 23 关于唯一性的断言指出,如此得到的友矩阵集合由 A 唯一确定. 于是我们证明了

定理 24　元素在域 F 中的任意矩阵 A, 在 F 上与一个且只与一个多项式

$$p_i(x)^{e_{i1}}, \quad \cdots, \quad p_i(x)^{e_{ir_i}}, \tag{40}$$

$$e_{i1}\geqslant\cdots\geqslant e_{ir_i}>0, \ i=1, \cdots, k$$

的友矩阵的直和相似, 这些多项式是首一不可约多项式 $p_1(x), \cdots, p_k(x)$ 的幂, A 的极小多项式是 $m(x)=p_1(x)^{e_{11}}p_2(x)^{e_{21}}\cdots p_k(x)^{e_{k1}}$.

这组多项式(40)是 A 在相似(在 F 上)之下的全系不变式, 称

为A的初等因子集合．把A表示为友矩阵的直和的表达式称为A的准素有理标准型（用"准素"这个词是因为用了不可约多项式的幂，用"有理"这个词是因为分析中只用到域F中的有理运算）．

推论 1 $n\times n$矩阵A的特征多项式是A的初等因子之积的$(-1)^n$倍．

证明 容易看出，矩阵$B_1, \cdots, B_q$的直和的特征多项式是$B_1, \cdots, B_q$的特征多项式的乘积．但是根据定理15，友矩阵C_f的特征多项式，除了符号外，就是$f(x)$．这两个事实同定理一起证明了推论1．

推论 2 方阵的特征值是它的极小多项式的根．

证明 因为极小多项式$m(x)$可整除特征多项式，所以极小多项式的任意根都是特征多项式的根，因此也是特征根（特征值）．反过来，根据推论1，特征多项式的任意根一定是某个初等因子$p_i(x)^{e_{ij}}$的根，因此根据定理，它是$m(x)$的根．

例 具有极小多项式$(x^2+1)(x+3)^2$的任意6×6有理矩阵与下列友矩阵直和之一相似：

$$C_{x^2+1}\oplus C_{x^2+1}\oplus C_{(x+3)^2},$$

$$C_{x^2+1}\oplus C_{(x+3)^2}\oplus C_{(x+3)^2},$$

$$C_{x^2+1}\oplus C_{(x+3)^2}\oplus C_{x+3}\oplus C_{x+3}.$$

第一种情形的特征多项式是$(x^2+1)^2(x+3)^2$；第二、三种情形的特征多项式都是$(x^2+1)(x+3)^4$．

在复数域上，首一不可约多项式只能是线性多项式$x-\lambda_i$，其中λ_i是标量．利用这个事实，对于复矩阵，或者更一般地，对于极小多项式是线性因子幂的乘积的任意矩阵，可以构造出不同的标准型．

这时，定理23中的每个T-循环子空间Z_a将有T-阶$(x-\lambda_i)^e$，其中λ_i为某个标量，e为正整数．对于Z_a的基底α，αT，$\cdots$，

αT^{e-1}，象定理 24 那样，T 可用$(x-\lambda_i)^e$ 的友矩阵来表示. 另一方面，考虑矢量 $\beta_1=\alpha, \beta_2=\alpha U, \cdots, \beta_e=\alpha U^{e-1}$，其中 $U=T-\lambda_i I$. 因为每个 β_j 是 αT^{j-1} 加上矢量 $\alpha T^k(k<j-1)$的某个线性组合，所以矢量 $\beta_1, \cdots, \beta_e$ 也是 Z_α 的一组基底. 为了得到T作用在 β_j 上的效果，注意

$$\beta_j T=\alpha U^{j-1}T=\alpha U^{j-1}(U+\lambda_i I)=\lambda_i \alpha U^{j-1}+\alpha U^j.$$

当 $j<e$ 时，这就得到 $\beta_j T=\lambda_i\beta_j+\beta_{j+1}$；当 $j=e$ 时，则有 $\alpha U^j=0$，$\beta_j T=\lambda_i\beta_j$. 现在对于这组基底，$T$ 可用矩阵

$$\begin{pmatrix} \lambda_i & 1 & & & 0 \\ & \lambda_i & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda_i & 1 \\ 0 & & & & \lambda_i \end{pmatrix}$$

来表示，它的行是 $\beta_j T$ 的坐标，上面列出的矩阵，主对角线上的元素都是 λ_i，紧靠主对角线上面的一条对角线上的元素都是 1，其他元素都是零. 称这样的矩阵为初等若当(Jordan)矩阵.

如果我们用上述类型的基底代替定理 24 中导出友矩阵的那组基底，则我们得到

定理 25 如果矩阵 A 的极小多项式在域 F 上分解成线性因子之积

$$m(x)=(x-\lambda_1)^{e_1}(x-\lambda_2)^{e_2}\cdots(x-\lambda_k)^{e_k}, \tag{41}$$

其中 $\lambda_1, \cdots, \lambda_k$ 是不同的，那么 A 在域 F 上与一个且只同一个初等若当矩阵的直和相似，这个直和至少包含一个属于特征根（特征值）λ_i 的 $e_i\times e_i$ 初等若当矩阵，并且没有更大的初等若当矩阵属于特征根(特征值) λ_i.

注意，在对角线上 λ_i 出现的个数是 λ_i 作为 A 的特征多项式的根的重数.

上面所得到的初等若当矩阵的直和，如果不计沿对角线排列的这些矩阵块的次序，它是唯一的，这种直和称为 A 的若当标准型．它可以应用到复数域上的任意矩阵．注意，若当标准型是由初等因子集合确定的，特别是，如果(41)中的所有的 e_i 都是 1，而且只有在这时，若当标准型成为对角矩阵，其对角线元素是 $\lambda_1, \cdots, \lambda_k$．于是，定理 22 的推论可作为上述定理一种特殊情形．

推论 任意复矩阵同若当标准型矩阵相似．

习 题

1．在有理数域上，对下列各矩阵求出所有可能的准素有理标准型：

(a) 5×5 矩阵，极小多项式是 $(x-1)^2$。

(b) 7×7 矩阵，极小多项式是 $(x^2-2)(x-1)$，特征多项式是 $(x^2-2)^2\cdot(x-1)^3$.

(c) 8×8 矩阵，极小多项式是 $(x^2+4)^2(x+8)^2$.

(d) 6×6 矩阵，特征多项式是 $(x^4-1)(x^2-1)$.

2．对具有下列各特征多项式的矩阵，列出所有可能的若当标准型：

(a) $(x-\lambda_1)^3(x-\lambda_2)^2$,

(b) $(x-\lambda_1)^5(x-\lambda_2)^3$,

(c) $(x-\lambda_1)(x-\lambda_2)^2(x-\lambda_3)^2$.

3．把正文中指出的初等若当矩阵表示成准素有理标准型．

4．(a) 证明：复矩阵和它的转置矩阵一定有相同的若当标准型．

(b) 推断它们总是相似的．

5．(a) 两个泡利(Pauli)“旋转矩阵”满足条件 $ST=-TS, S^2=T^2=I$，并且是埃尔米特矩阵．证明：$U=iST$ 是埃尔米特矩阵，并满足 $TU=-UT, U^2=I$.

(b) 证明：如果 S 是 2×2 矩阵，则 S 与 $\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}$ 相似，并且对于这组坐标，$T=\begin{pmatrix}0 & b\\ b^{-1} & 0\end{pmatrix}$，其中 b 是某数．

*6．用 §10.9 的方法证明：任意线性变换 $T:V\to V$ 把 V 分解成具有 T-阶 $f_1(x), \cdots, f_r(x)$ 的 T-循环子空间的直和，这里 $f_i(x)\mid f_{i-1}(x)$ $(i=2, \cdots, r)$，而且 $f_1(x)$ 是 T 的极小多项式．

第十一章 布尔代数与格

§11.1 基本定义

我们现在将从近世代数的观点更严密地分析“集合”(或类)和“子集合”的基本概念，这些概念在§1.11中已有过简短介绍．假设 I 为任意集合，而 X, Y, Z 表示 I 的子集．比如 I 是正方形，X, Y, Z 三个是位于 I 中的全等的互相交叠的圆形，如图1的“维恩(Venn)图”所示．

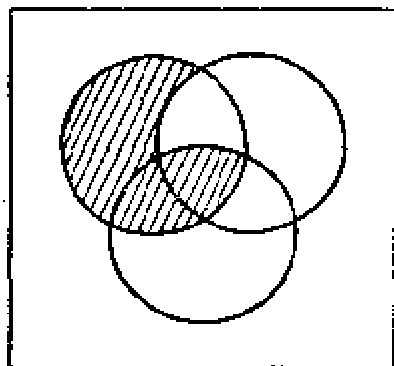


图 1

当 X 是 Y 的子集，即 X 的每个元素都在 Y 中，我们记作 $X\subset Y$(或 $Y\supset X$)．这个关系也可称为 X“包含”在 Y 中．

包含关系满足自反律：这是显然的，因为任意集合 X 是它本身的子集．包含关系也满足传递律：因为如果 X 的每个元素在 Y 中，并且 Y 的每个元素在 Z 中，那么显然 X 的每个元素在 Z 中．但是，包含关系不满足对称律．反之，如果 $X\subset Y$ 且 $Y\subset X$，那么 X 和 Y 一定包含同样的元素，因此 $X=Y$．

概括起来，集合的包含关系与算术的不等关系都具有下列性质：

自反律　对一切 X，有 $X\subset X$．

反对称律　如果 $X\subset Y$ 且 $Y\subset X$，那么 $X=Y$．

传递律　如果 $X\subset Y$ 且 $Y\subset Z$，那么 $X\subset Z$．

但是，“对于给定的两个集合 X 和 Y，不是 $X\subset Y$，就是 $Y\subset X$”这个

命题是不正确的.

因此两个集合 X 和 Y 的包含关系有四种可能的方式. 一种可能是 $X \subset Y$ 并且 $Y \subset X$, 在这种情况下, 根据反对称律有 $X=Y$. 另一种可能是 $X \subset Y$ 但不满足 $Y \subset X$, 在这种情况下, 我们称 X 真包含在 Y 中, 并记作 $X<Y$ 或 $Y>X$. 我们还可以有 $Y \subset X$ 但不满足 $X \subset Y$, 在这种情况下, 说 X 真包含 Y. 最后, 我们有既不是 $X \subset Y$, 也不是 $Y \subset X$, 在这种情况下称 X 和 Y 是不可比的. 由于不可比集合的存在, 才使包含关系不同于实数间的不等关系.

已知集合 I 的子集中不仅有包含关系, 而且可以通过两种二元运算"并"与"交"把它们联系起来, 这两种运算类似于普通的"加"与"乘". 这种类比的程度和重要性首先是由英国数学家布尔(George Boole, 1815—1864)发现的, 他在一百多年以前就建立了集合代数的理论.

我们把 X 和 Y 的交 (记作 $X \cap Y$) 定义为既在 X 中又在 Y 中的所有元素的集合, 把 X 和 Y 的并(记作 $X \cup Y$)定义为或者在 X 中或者在 Y 中或者同时在两个集合之中的所有元素的集合. 符号"$\cap$"和"$\cup$"分别称为"求交"运算和"求并"运算.

最后, 我们用 X'(读作"X 的补")表示不在 X 中的所有元素的集合. 例如, I' 是空集 $\varnothing$, 它不包含任何元素. 这是因为我们所考虑的只是 I 的子集.

集合的代数运算可以通过图 1 的维恩图加以说明. 在这个图中, X, Y, Z 是三个交叠圆形的内部, 这些区域在正方形 I 中的组合可以用适当的阴影区域来表示. 例如, Y' 是 Y 的外部, $X \cap (Y' \cup Z)$ 是图中的阴影区域.

习　题

1. X、Y、Z 的维恩图把正方形分割成八个不交叠的区域. 用 X, Y,

Z 的代数组合注明每个这样的区域.

2. 在维恩图上把下列各区域画上阴影:

$(X' \cap Y) \cup (X \cap Z')$, $(X \cup Y)' \cap Z$,

$(X \cup Y') \cup Z'$.

3. 通过对维恩图上适当的区域画出阴影, 确定下列方程中哪些是成立的:

(a) $(X' \cup Y)' = X \cap Y'$,

(b) $X' \cup Y' = (X \cup Y)'$,

(c) $(X \cup Y) \cap Z = (X \cup Z) \cap Y$,

(d) $X \cup (Y \cap Z)' = (X \cup Y') \cap Z'$.

§11.2 定律: 同算术定律类比

我们现在略为详细地描述一下集合代数与普通算术之间的类似, 并用来定义布尔代数. “$\cap$, $\cup$”和普通的“$\cdot$, $+$”之间的类似, 由下列定律作部分的描述, 这些定律的正确性是显然的.

幂等律 $X \cap X = X$ 和 $X \cup X = X$.

交换律 $X \cap Y = Y \cap X$ 和 $X \cup Y = Y \cup X$.

结合律 $X \cap (Y \cap Z) = (X \cap Y) \cap Z$

和 $X \cup (Y \cup Z) = (X \cup Y) \cup Z$.

分配律 $X \cap (Y \cup Z) = (X \cap Y) \cup (X \cap Z)$

和 $X \cup (Y \cap Z) = (X \cup Y) \cap (X \cup Z)$.

显然, 除了幂等律和第二分配律之外, 所有这些定律都与大家所熟悉的“$\cdot$”与“$+$”的性质相对应, 这些性质在第一章已作为公设给出.

下面的基本定律把交和并相互联系起来, 而且把交、并和包含联系起来.

相容律 $X \subset Y$, $X \cap Y = X$ 和 $X \cup Y = Y$ 这三个条件是互相等价的.

还有，空集用$\varnothing$表示，$\varnothing$和I具有下列特殊性质：

泛界　$\varnothing \subset X \subset I$，对一切$X$.

　交　$\varnothing \cap X = \varnothing$ 和 $I \cap X = X$.

　并　$\varnothing \cup X = X$ 和 $I \cup X = I$.

前三个交和并的性质与普通算术中 0 和 1 的性质相类似.

最后，下面三个新的定律把交、并和补联系起来.

互补律　$X \cap X' = \varnothing$ 和 $X \cup X' = I$.

对偶律　$(X \cap Y)' = X' \cup Y'$ 和 $(X \cup Y)' = X' \cap Y'$.

对合律　$(X')' = X$.

如果把X'解释为$1-X$，并假定$XX=X$，那么互补律和对合律与普通算术定律相对应.

上述定律可以用各种方法证明．第一，我们可用特殊例子通过“归纳推理”来检验它们．维恩图提供了一个合适的例子．如果X和Y分别是图 2 中左圆形和右圆形的内部，那么对于区域X'画出水平直线的阴影，对于区域Y'画出垂直直线的阴影．那么十字阴影线的区域就是$X' \cap Y'$．由右图立即看出，这个区域是并$X \cup Y$的补．这就是第二对偶律所描述的．就我们的常识而言，可以承认这样的论证，但是，数学上这是不允许的，因为在数学推理中，只允许演绎证明.

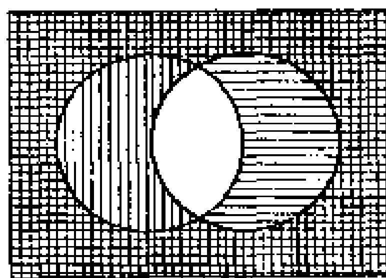


图　2

第二，我们可以把I的元素分为四种可能情况来考虑：(i)元素既在X中又在Y中；(ii)元素在X中但不在Y中；(iii)元素在Y中但不在X中；(iv)元素既不在X中也不在Y中．比如，(i)类元素在$X \cap Y$中因此不在$(X \cap Y)'$中，不在$X' \cup Y'$中；而(ii)类元素在$(X \cap Y)'$和Y'中，因此在$X' \cup Y'$中．再看其他两类元素，也是

这个情况. 因此我们看出, $(X\cap Y)'$ 和 $X'\cup Y'$ 具有相同的元素, 这就是第一对偶律. 注意, 对于两个集合 X 和 Y, 元素的四种可能情况用上面这个维恩图的四个区域中的点来表示; 而对于三个集合, 元素有八种情况, 对应于八个区域(图 1).

第三, 我们可以用"求并"和"求交"运算的叙述性的定义重述这些定律. 例如, 考虑分配律. 这里

"b 在 $X\cap(Y\cup Z)$ 中"是指"b 既在 X 中又在 Y 或 Z 中",

"b 在 $(X\cap Y)\cup(X\cap Z)$ 中" 是指 "b 或者既在 X 中又在 Y 中, 或者既在 X 中又在 Z 中".

按照连词"既…, 又…"及"或者…, 或者…"的通常用法, 稍微"翻译" 一下就使我们确信这两种叙述是等价的. 分配律的这个证明表明, 集合代数中的定律怎样翻译为"既…, 又…"、"或"、"非"这些词的性质. 如果我们假定这些性质是基本的, 那么象我们进行平常的数学推理那样, 就能从这些性质证明上述所有关于集合的定律.

习　　题

1. 用维恩图验证分配律.

2. 用细分几种情况的方法验证结合律、交换律和相容律.

3. 利用上述第三种方法中的"既…, 又…"、"或…, 或…"、"非"等词重述互补律、对偶律和对合律.

4. (a) 在处理四个集合的代数表达式时, 考虑它们元素的所有可能情况, 会出现多少种情况?

*(b) 画出一个四集合图, 它把元素的每一种可能情况表示成一个区域.

(c) 证明: 不存在这样的图形, 在这个图中四个给定的集合都是圆形.

5. 证明: $\varnothing$ 和 I 的交和并的性质可以从泛界性和相容律推导出.

6. 在相容律中, 用"必要性"和"充分性"代替 "等价" 而得到六个推断. 证明: 有三个推断对于满足 $0\leqslant x,\ y\leqslant 1$ 的实数 x, y 是成立的.

7. 在习题 6 中，如果$\varnothing$用数 0 代替，I 用数 1 代替，那么$\varnothing$，I 的交和并的性质中哪些是不成立的？如果 X'用 $1-x$ 代替，Y'用 $1-y$ 代替，那么关于补的性质中哪些是不成立的？

8. 证明：对 I 的任意子集 X, Y，$X\subset Y$ 当且仅当 $X'\cup Y=I$.

§11.3 布尔代数

我们不再关心由基本逻辑法则来推导上述代数定律，而是把这些定律中最基本的定律作为公设(象第一章中的算术定律那样)，然后再从这些公设导出尽可能多的有意义的结论来．因此，我们现在用稍微不同的记号给出基本定义，用这些记号是为了强调这些公设可以用于不同于集合的其他对象．

定义 具有下列性质的元素 a, b, c, … 的集合 B 称为布尔代数：

(i) B 有两个二元运算$\wedge$(楔形)和$\vee$(v 形)，它们满足

幂等律　$a\wedge a=a\vee a=a$.

交换律　$a\wedge b=b\wedge a$, $a\vee b=b\vee a$.

结合律　$a\wedge(b\wedge c)=(a\wedge b)\wedge c$,

$$a\vee(b\vee c)=(a\vee b)\vee c.$$

(ii) 这两个运算满足吸收律

$$a\wedge(a\vee b)=a\vee(a\wedge b)=a.$$

(iii) 这两个运算是互相可分配的

$$a\wedge(b\vee c)=(a\wedge b)\vee(a\wedge c),$$

$$a\vee(b\wedge c)=(a\vee b)\wedge(a\vee c).$$

(iv) B 包含泛界 O, I，它们满足

$$O\wedge a=O,\quad O\vee a=a,$$

$$I\wedge a=a,\quad I\vee a=I.$$

(v) B 有求补的一元运算 $a\longmapsto a'$，它遵循下面的互补律

$$a\wedge a'=O,\quad a\vee a'=I.$$

当然上述所有定律都假定对一切 $a,b,c\in B$ 是成立的.

利用这个定义，§11.1和§11.2的结论可以概括为下面的命题.

定理1 在交、并和补三种运算之下，任意集合 I 的全体子集构成一个布尔代数.

为了更有选择地说明上述这些公设的意义，我们现在描述几个例子，在这些例子中，有一些公设成立，但不全都成立.

例1 设 L 是以 n 维欧几里得矢量空间（§7.10）的子空间为元素的集合. 这里定义 $S\wedge T=S\cap T$ 是 S 和 T 的交，$S\vee T=S+T$ 是 S 和 T 的线性和，O 是零矢量 $\mathbf{0}$，I 是整个空间，S' 是子空间 S 的正交补空间 $S^{\perp}$.

那么，公设(i)，(ii)，(iv)和(v)都满足. 但是分配律(iii)不满足.（例如，设 S,T,U 分别是平面上由(1,0)，(0,1)，(1,1)张成的子空间.）

例2 设 L 是以有限群 G 的正规子群 M，N，…为元素的集合. 设 $M\wedge N=M\cap N$ 是 M 和 N 的交，而 $M\vee N=MN$ 是由所有乘积 $xy(x\in M,\ y\in N)$ 组成的集合. 那么 $M\wedge N$ 和 $M\vee N$ 都是 G 的正规子群. 如果 O 表示群单位元素1，I 是群 G 本身，那么虽然(iii)和(v)一般都不满足，但公设(i)，(ii)和(iv)都满足.

因为例1和例2中构造的系统都满足公设(i)和(ii)，所以它们在下述意义下是格.

定义 如果集合 L 有两个二元运算①$\wedge$和$\vee$，它们满足幂等律、交换律和结合律，并且满足吸收律(ii)，那么 L 是一个格. 如果除此之外还满足分配律(iii)，那么 L 称为分配格.

①$\wedge$运算也称为交(*meet*)，$\vee$运算也称为并(*join*)，我们将交替使用这些名称.

例如，如果所有多边形区域的集合L包含空集$\varnothing$，并且面积为零的集合可以忽略不计，那么集合L在交和并运算之下是一个分配格．再如，在所有正整数的集合Z^+中，如果定义$m\wedge n$是m和n的最大公因子，$m\vee n$是m和n的最小公倍数，那么Z^+是一个分配格．

上面作为公设的各种定律有很多有趣的代数结论，我们现在来推导其中最简单的几个．

结合律和交换律的作用已经在§1.5中研究过了．结合律实际上意味着我们可以不用括号来组成多重交或多重并；交换律意味着，在只含有$\vee$或只含有$\wedge$的表达式中，各项可以按照我们喜欢的任何方式排列．

连同上面的定律，幂等律的作用显然是允许我们消去重复出现的项——只留下一个已知项，其余重复出现的项全都消去．概括起来我们有

引理1 设f和g是由符号$\vee$和所有字母$a_1, \cdots, a_n$（可能有些字母重复）构成的两个表达式，那么由幂等律、交换律和结合律可推出$f=g$．对于只含有$\wedge$的表达式，上述结论同样成立．

设N是下标$i=1, 2, \cdots, n$的集合，我们可以不含糊地用

$$\bigvee_N a_i \text{ 或 } \bigvee_{i=1}^n a_i$$

和

$$\bigwedge_N a_i \text{ 或 } \bigwedge_{i=1}^n a_i$$

分别表示所有a_i的并(join)和交(meet)．这些记号类似于代数记号Σ和Π．

再有，我们从交换律、结合律和分配律出发，可以用归纳法象§1.5那样导出一般分配律如下

$$x\wedge(y_1\vee\cdots\vee y_n)=(x\wedge y_1)\vee\cdots\vee(x\wedge y_n),$$
$$x\vee(y_1\wedge\cdots\wedge y_n)=(x\vee y_1)\wedge\cdots\wedge(x\vee y_n),$$
$$(x_1\vee\cdots\vee x_m)\wedge(y_1\vee\cdots\vee y_n)$$
$$=(x_1\wedge y_1)\vee(x_1\wedge y_2)\vee\cdots\vee(x_m\wedge y_n).$$

习　　题

1. 用归纳法详细证明：

(a) 在任意分配格中，$x\wedge\bigvee_{i=1}^{n}y_i=\bigvee_{i=1}^{n}(x\wedge y_i)$.

(b) 在任意布尔代数中，$\left(\bigvee_{i=1}^{n}x_i\right)'=\bigwedge_{i=1}^{n}x_i'$.

*2. 用归纳法详细证明：在任意分配格中，

$$\left(\bigvee_{i=1}^{m}x_i\right)\wedge\left(\bigvee_{j=1}^{n}y_j\right)=\bigvee_{i=1}^{m}\left[\bigvee_{j=1}^{n}(x_i\wedge y_j)\right].$$

3. 详细证明：当 $n>1$ 时，例 1 定义了一个格，但不是分配格.

4. 证明：当 G 为循环群时，例 2 定义了一个分配格.

5. 证明：四元素群的所有子群组成的格不是分配格.

§11.4　其他基本定律的推导

我们指出，上面列出的关于布尔代数的公设可推出 §11.1 和 §11.2 中讨论的集合代数的其他基本公式．例如，它们可推出 O 和 I 的唯一性，这些我们并没有假定过.

引理 2　*在任意布尔代数中，恒等式 $a\wedge x=a$ 和 $a\vee x=x$（对一切 x）中每一个都可推出 $a=O$．对偶地有，恒等式 $a\vee x=a$ 和 $a\wedge x=x$（对一切 x）中每一个都可推出 $a=I$.*

证明　如果对所有 x，$a\wedge x=a$，那么特别有 $a\wedge O=a$；但是由(iv)有 $a\wedge O=O$，因此 $a=O$．同样，如果对所有 x，$a\vee x=x$，那么 $a\vee O=O$；但是由(iv)有 $a\vee O=a$，因此又有 $a=O$．I 的唯一性的

证明类似.

引理 3 对任意格中的元素 a, b, $a\wedge b=a$ 成立当且仅当 $a\vee b=b$.

证明 如果 $a\vee b=b$, 那么根据吸收律(ii), 有 $a\wedge b=a\wedge(a\vee b)=a$. 反之, 如果 $a\wedge b=a$, 则 $a\vee b=(a\wedge b)\vee b$. 因此根据交换律, $a\vee b=b\vee(b\wedge a)=b$, 这里最后一步又用到(ii).

推论 在布尔代数的定义中, 条件(iv)可由下列公设中的任何一个来代替:

(iv′) 对一切 x, 有 $x\wedge O=O$ 和 $x\vee I=I$.

(iv″) 对一切 x, 有 $O\vee x=x$ 和 $I\wedge x=x$.

上面给出的布尔代数的定义没有提到包含关系, 即使包含关系是所有概念中最基本的. 我们现在来定义这个关系, 并由上述公设推导它的基本性质. 其证明重述相容律, 相容律的一部分已经证过了, 如上面引理 3 所述.

定义 定义 $a\leqslant b$ 是指 $a\wedge b=a$, 或者指 $a\vee b=b$(根据引理3, 这两个说法是等价的).

引理 4 在任意格中, 关系 $a\leqslant b$ 满足自反律、反对称律和传递律.

证明 因为 $a\wedge a=a$, 所以对一切 a, 有 $a\leqslant a$, 这就证明了自反律. 再有, 由 $a\leqslant b$ 和 $b\leqslant a$ 可推出

$$a=a\wedge b=b\wedge a=b,$$

这就证明了反对称律. 最后, 由 $a\leqslant b$ 和 $b\leqslant c$ 推出 $a=a\wedge b=a\wedge(b\wedge c)=(a\wedge b)\wedge c=a\wedge c$, 因此 $a\leqslant c$. 这就证明了传递律. 证毕

吸收律的作用在上面引理 2 和引理 3 的证明中已经显示出来. 幂等律在格的定义中是多余的, 实际上, 由吸收律、交换律和结合律可推出幂等律: 因为吸收律是说, 对所有 x, z, 有 $x=x\wedge(x$

$\vee z$). 设 $z=x\wedge y$, 我们推出, 对所有 x, y, 有$x=x\wedge[x\vee(x\wedge y)]$; 再应用对偶的吸收律 $x\vee(x\wedge y)=x$, 于是我们就得到 $x=x\wedge x$(这就是幂等律). $x=x\vee x$ 的证明类似, 只须把$\wedge$和$\vee$互换.

引理 5 在任意分配格中, 由 $a\vee x=a\vee y$ 和 $a\wedge x=a\wedge y$ 一起可推出 $x=y$.

证明 通过等式替换, 并逐次应用吸收律和分配律, 我们有

$$
\begin{aligned}
x&=x\wedge(x\vee a)=x\wedge(y\vee a)\\
&=(x\wedge y)\vee(x\wedge a)=(y\wedge x)\vee(y\wedge a)\\
&=y\wedge(x\vee a)=y\wedge(y\vee a)=y.
\end{aligned}
$$

现在我们回想一下求补运算 $a\longmapsto a'$ 满足

$$a\wedge a'=O \text{ 和 } a\vee a'=I.$$

但是任意满足 $a\wedge x=O$ 和 $a\vee x=I$ 的元素 x, 根据引理 5, 它一定满足 $x=a'$. 换句话说, 补 a' 由布尔代数定义中的互补律(v)唯一确定. 我们现在证明补集的其余性质(对偶律和对合律)在任意布尔代数中也都成立.

引理 6 在任意布尔代数中, 我们有

$$(x')'=x,\ (x\wedge y)'=x'\vee y' \text{ 和} (x\vee y)'=x'\wedge y'. \tag{1}$$

证明 "x' 是 x 的补"这个说法由交换律可推出"x是x' 的补", 这因为 $x'\wedge x=x\wedge x'=O$ 和 $x'\vee x=x\vee x'=I$. 但是我们刚刚证明过补是唯一的, 因此 x 是 x' 的唯一的补, 于是$(x')'=x$. 再有, 根据分配律有

$$
\begin{aligned}
(x\wedge y)\wedge(x'\vee y')&=(x\wedge y\wedge x')\vee(x\wedge y\wedge y')\\
&=[(x\wedge x')\wedge y]\vee(x\wedge O)\\
&=[O\wedge y]\vee O=O\vee O=O.\\
(x\wedge y)\vee(x'\vee y')&=(x\vee x'\vee y')\wedge(y\vee x'\vee y')\\
&=(I\vee y')\wedge(y\vee y'\vee x')\\
&=I\wedge(I\vee x')=I.
\end{aligned}
$$

这就证明了 $x'\vee y'$ 是 $x\wedge y$ 的补. 因此,再根据补的唯一性,$x'\vee y'=(x\wedge y)'$ 是 $x\wedge y$ 的补. 恒等式 $(x\vee y)'=x'\wedge y'$ 可以类似地证明.

推论 为了求出由带撇和不带撇的字母通过多重$\vee$和$\wedge$(但不用加撇的括号)构成的表达式的补,可以把表达式中的$\vee$和$\wedge$全都互换,并把每个不带撇的字母加上撇,把每个带撇字母的撇去掉.

例如,根据这个法则,$(x'\wedge y)\vee(z\wedge w')$ 的补是 $(x\vee y')\wedge(z'\vee w)$.

证明 如果在已知表达式 f 中字母的个数n(重复的也计算在内)是 1,那么推论是正确的,这因为$(x)'=x'$,$(x')'=x$. 如果不然,因为表达式中的括号都不带撇,所以我们可以把它写成 $f=a\wedge b$ 或 $f=a\vee b$,由此分别得到 $f'=a'\vee b'$ 或 $f'=a'\wedge b'$. 但是表达式 a 和 b 包含的字母比 f 包含的字母少,因此,对 n 用归纳法,我们可以假定推论对于 a 和 b 都是正确的. 再代入表达式 $f'=a'\vee b'$ 或 $f'=a'\wedge b'$ 中,我们就得到所要求的补的公式.

习 题

1. 证明: 幂等律 $x\vee x=x$ 可由交换律、结合律和吸收律推出.

习题 2～10 是在布尔代数的条件之下.

2. 详细证明: $(x\vee y)'=x'\wedge y'$.
3. 化简下列布尔表达式:

   (a) $(x'\wedge y')'$,

   (b) $(a\vee b)\vee(c\vee a)\vee(b\vee c)$,

   (c) $(x\wedge y)\vee(z\wedge x)\vee(x'\vee y')'$.
4. 证明: $(x\wedge y)\vee(x\wedge y')\vee(x'\wedge y)\vee(x'\wedge y')=I$. 利用两个圆的维恩图加以解释.
5. 证明: $x=y$ 当且仅当$(x\wedge y')\vee(x'\wedge y)=O$.

6. 证明包拉茨基（Poretzky）定律：已知 x 和 t，$x=O$ 当且仅当 $t=(x\wedge t')\vee(x'\wedge t)$.

7. (a) 证明：$y\leqslant x'$ 当且仅当 $x\wedge y=O$.

(b) 证明：$y\geqslant x'$ 当且仅当 $x\vee y=I$.

8. 求出下列表达式的补：

(a) $x\vee y\vee z'$,

(b) $(x\vee y'\vee z')\wedge(x\vee(y\vee z'))$,

(c) $x\vee(y\wedge(z\vee w'))$,

(d) $(x'\vee y)'\wedge(x\vee y')$,

9. 把引理 6 的推论的论证方法应用到表达式 $(x'\wedge y\wedge z')\vee(x\wedge y')$ 上，并说明每一步的理由.

10. 证明：$(x\vee y)\wedge(x'\vee z)=(x'\wedge y)\vee(x\wedge z)$.

11. 证明：在任意分配格中，有

$$(x\wedge y)\vee(y\wedge z)\vee(z\wedge x)=(x\vee y)\wedge(y\vee z)\wedge(z\vee x).$$

12. 具有泛界 O, I 的格 L 中的元素 a，如果对某个 $x\in L$，有 $a\wedge x=O$ 和 $a\vee x=I$，那么称 a 为有补元素. 证明：如果 a 和 b 都是分配格的有补元素，那么 $a\wedge b$ 和 $a\vee b$ 也都是有补元素.

§11.5 布尔多项式的标准型

在前一节里，我们已经研究了由 $\wedge$, $\vee$ 和 $'$ 运算构成的各种表达式. 这样的表达式称为“布尔多项式”（或“布尔函数”），显然，它类似于普通多项式.

我们现在来定义布尔代数 B 的子代数为 B 中这样的非空子集 S：如果它包含任意两个元素 x 和 y，那么它也包含 $x\wedge y$, $x\vee y$, x'（因而也包含 $O=x\wedge x'$ 和 I）. 给定 B 的一个任意非空子集 X，那么所有值 $p(x_1,\cdots,x_n)$（元素 $x_i\in X$）组成的集合显然是 B 的包含 X 的最小子代数. 同群的情形一样，称这个子代数是由 X 生成的. 例如，任意一个元素 x 生成的子代数由 x, x', O, I 四个元素组成.

这是下面使人感到惊奇的事实的一个特殊情形：这个事实是

n 个变量 $x_1, \cdots, x_n$ 的不同布尔多项式的个数等于 2^{2^n}. 现在我们来证明它, 以多项式

$$f(x, y, z) = [x \vee z \vee (y \vee z)']' \vee (y \wedge x)$$

为例进行论证.

第一, 如果多项式中任何括号的外边出现撇, 那么总可以应用对偶律(如 §11.4 的引理 6)把它移到括号里边. 当所有的撇都移到括号的最里边时, 多项式变成只含有带撇字母和不带撇字母以及作用在它们上面的$\vee$和$\wedge$的表达式. 例如上述例子中,

$$f = [x' \wedge z' \wedge (y \vee z)] \vee (y \wedge x).$$

第二, 如果任意$\wedge$在括号外边, 而括号里包含$\vee$, 那么根据分配律, $\wedge$可以移到括号里边, 象$c \wedge (a \vee b) = (c \wedge a) \vee (c \wedge b)$那样. 结果得到一个多项式, 其中所有的交$\wedge$先组合起来, 然后再按并$\vee$组成, 也就是说, 这个表达式是某些项 $T_1, \cdots, T_k$ 的并, 其中每个 $T_i (i = 1, 2, \cdots, k)$是一些带撇和不带撇字母的交. 在上面例子中,

$$f = (x' \wedge z' \wedge y) \vee (x' \wedge z' \wedge z) \vee (y \wedge x).$$

第三, 某些表达式可以缩短或省略去. 如果字母"c"在一项中出现两次, 则可略去一个"c", 这因为$c \wedge c = c$. 如果带撇的 c 和不带撇的 c 同时出现在由$\wedge$连接的项中, 那么整个项是 O, 这因为对一切 a, 有 $c \wedge a \wedge c' = O$; 因此这一项在由$\vee$连接的项中可以略去, 因为对一切 b, 有 $O \vee b = b$. 例如上面的例子中

$$f = (x' \wedge z' \wedge y) \vee (y \wedge x).$$

现在, 如果某一项 T_k 不包含字母 c, 我们可以写成

$$T_k = T_k \wedge I = T_k \wedge (c \vee c') = (T_k \wedge c) \vee (T_k \wedge c'),$$

这里是用两项代替 T_k, 每项中 c 恰好出现一次. 例如在我们的例子中,

$$f = (x' \wedge z' \wedge y) \vee (y \wedge x \wedge z) \vee (y \wedge x \wedge z').$$

最后, 每一项中出现的字母可以重新排列, 使得它们按自然顺

序出现，例如

$$f=(x'\wedge y\wedge z')\vee(x\wedge y\wedge z)\vee(x\wedge y\wedge z')$$

这称为 f 的析取标准型；于是我们就证明了下面的引理.

引理 任意 $x_1, \cdots, x_n$ 的布尔多项式可以或者化为 O，或者化为某些项 T_k 的并，其中 T_k 具有形式

$$T_k=q_1\wedge q_2\wedge\cdots\wedge q_n \qquad (\text{每个 } q_j=x_j \text{ 或 } x'_j). \tag{2}$$

也就是说，可以化为析取标准型.

因为每个 q_j 都有两种可能，所以我们看到，T_k 恰有 2^n 种可能. 例如，当 $n=3$ 时，任意布尔多项式用我们的方法可以化为 O 或者化为项

$$\begin{array}{llll} x\wedge y\wedge z, & x'\wedge y\wedge z, & x\wedge y'\wedge z, & x\wedge y\wedge z', \\ x\wedge y'\wedge z', & x'\wedge y\wedge z', & x'\wedge y'\wedge z, & x'\wedge y'\wedge z' \end{array} \tag{3}$$

的某一个并. 图 1 的三个圆把正方形分成八个区域，这八个多项式就表示这八个区域，这个事实并非偶然. 这在几何上意味着，三个圆 X, Y, Z 的任何布尔组合是图中八个区域的某种选择的并.

象(3)中列举的那些基本项称为极小布尔多项式. 换句话说，n 个变量 $x_1, \cdots, x_n$ 的极小布尔多项式 $M(x_1, \cdots, x_n)$ 是 n 个元素的交 $\bigwedge\limits_{i=1}^{n} q_i$，其中第 i 个元素 q_i 或者是 x_i 或者是 x'_i. 于是我们就证明了

定理 2 任意已知的 $x_1, \cdots, x_n$ 的布尔多项式或者等于 O 或者等于一组极小多项式(记为 S)的并.

现在赋给每个 M 一个 n 位二进数 $\eta(M)=y_1y_2\cdots y_n$，这里数字 y_i 是 1 或 0 应根据上面的 $M=\bigwedge\limits_{i=1}^{n} q_i$ 中 q_i 是 x_i 或 x'_i 而定. 那么函数 η: $M\longmapsto\eta(M)$ 是由 $x_1, \cdots, x_n$ 的极小多项式的集合到所有

2^n 个 n 位二进数的集合 I 的双射. 例如在(3)中,这些极小多项式所对应的 $\eta(M)$ 值是

$$111,\ 011,\ 101,\ 110, 100,\ 010,\ 001,\ 000.$$

另一方面,$\eta(M)$ 可以看作矢量 $\eta=(y_1, y_2, \cdots, y_n)\in \mathbf{Z}_2^n$, 并且可以认为每个布尔多项式 $\bigvee_S M_\eta(x_1, \cdots, x_n)$ 对应于这些矢量组成的集合.

如果 $S_i\subset I$ 是由那些第 i 位数字 $y_i=1$ 的二进数组成, 那么 S_i' 将由那些 $y_i=0$ 的二进数组成. 因此(参看下面习题 9), 表示已知极小多项式 $M(S_1, \cdots, S_n)$ 的集合是由单个二进数 $\alpha(M)=a_1a_2\cdots a_n$ 组成,它的第 i 个数字 a_i 是 0 还是 1, 视 S_i 在 M 中带撇还是不带撇而定. 显然,不同的极小多项式 $M=M_\alpha$ 要用不同的二进数来表示, 因此,M_α 的不同集合中极小多项式的并表示 I 的不同子集. 这就证明了下面的结果.

推论 恰有 2^{2^n} 个不同的 n 变量的布尔函数.

我们现在可以用系统的方法来代替布尔多项式偶而的运算. 任意给出的布尔代数中的方程 $E_1=E_2$, 只要把两边都化为析取标准型就可确定这个方程是正确还是错误.

习 题

1. 把下列各表达式化为标准型:

   (a) $(x\vee y)\wedge(z'\wedge y)'$;

   (b) $(x\vee y)\wedge(y\vee z)\wedge(x\vee z)$.

2. 用把等式两边化为(析取)标准型的方法检验下列给出的各方程的正确性:

   (a) $[x\wedge(y\vee z)']'=(x\wedge y)'\vee(x\wedge z)$,

   (b) $x=(x'\vee y')'\vee[z\vee(x\vee y)']$

3. 证明: 每个布尔多项式具有对偶标准型, 它是某些"素多项式"的

"交". 详细描述这些素多项式, 并指出它们是极小多项式的补. 这个结果与多项式分解定理"域上的每个普通多项式可唯一地表示成不可约多项式的乘积"有什么类似之处?

4. 用习题 3 的标准型检验习题 2(a) 的方程.

5. 证明: $f(x, y)$ 的标准型是

$$f(x,y)=[f(I,I)\wedge x\wedge y]\vee[f(I,O)\wedge x\wedge y']$$
$$\vee[f(O,I)\wedge x'\wedge y]\vee[f(O,O)\wedge x'\wedge y'].$$

6. 证明: 任意两个不同的极小多项式的交是 O.

7. 根据一般分配律展开 $I=(x_1\vee x_1')\wedge\cdots\wedge(x_n\vee x_n')$, 来证明 I 是所有极小布尔多项式的集合的并.

8. 由习题 7 和 $x_i=x_i\wedge I$ 来证明: 每个 x_i 是所有那些第 i 项为 x_i 的极小多项式的并.

9. (a) 设 $\bigvee_A M_\alpha$ 表示集合 A 中的所有极小多项式的并, 证明:

$$\left(\bigvee_A M_\alpha\right)\vee\left(\bigvee_B M_\beta\right)=\bigvee_{A\cup B} M_\gamma$$

$$\left(\bigvee_A M_\alpha\right)\wedge\left(\bigvee_B M_\beta\right)=\bigvee_{A\cap B} M_\gamma.$$

(b) 证明: 如果我们定义极小多项式的空集的并 $\bigvee_\phi M_\alpha$ 是 O, 那么上述公式仍然成立.

*10. 利用习题 7 和习题 9 证明: $\left(\bigvee_A M_\alpha\right)'=\bigvee_{A'} M_\alpha$. (提示: 运用 §11.4 引理 6.)

*11. 只用习题 8~10 独立地证明: 每个布尔多项式可以写成极小多项式的并.

§11.6 半 序

前面很少用到"包含"的自反律、反对称律和传递律, 然而这些定律是所有定律中最基本的, 因此可以把它们应用到许多非布尔代数的系统中去.

例如，对于一个集合的所有子集组成的系统，这些定律显然成立，这些子集是按照任意特殊性质来划分的（记作$\subset$或$\leqslant$）。例如，这些定律对于任意群的全体子群（或者全体正规子群）。对于任意域的全体子域，对于任意线性空间的全体子空间，等等，都成立——即使所有这些系统都不构成布尔代数。这些定律对于实数之间的“小于或等于”关系 $x\leqslant y$，对于正整数之间的整除关系 $x|y$，等等，也都成立。

这些例子暗示了“半序”这个抽象概念。“半序”是指任意满足自反律、反对称律和传递律的关系。

定义 *一个具有二元关系$\leqslant$的集合 P，如果这个关系满足自反律、反对称律和传递律，那么称P为半序集。*

对于这种类型的任意关系 $a\leqslant b$（读作“b包含a”），我们可以定义 $a<b$ 的意思是：$a\leqslant b$ 但 $a\neq b$；而当 $a<b$，并且没有 x 能满足 $a<x<b$，这时可称 b *覆盖* a。

下面引理指出，任意格可看作一个半序集（它的完整含义将在下一节里说明）。

引理 *在任意格中，如果关系 $x\leqslant y$ 的意思是 $x\wedge y=x$（等价于 $x\vee y=y$），那么这个关系是一个半序。*

具有有限个元素的半序集可以用图方便地表示出来。系统中的每个元素用小圆圈表示，如果 $a>b$，则对应 a 的小圆圈画在对应 b 的小圆圈之上。然后对于 a 覆盖 b 的情形，我们从 a 到 b 画一条下降的直线。我们可以从图上重新构造关系 $a\geqslant b$，因为 $a>b$ 当且仅当在图中从 b 出发沿着某些上升的直线段爬到 a。

例如，在图 3 中，第一图表示四元素群的所有子群组成的系统；第二图表示三点集的所有子集组成的布尔代数；第三图表示数 1, 2, 4, 8 在整除关系之下组成的系统。其他几个是随便构造的，它告诉我们怎样只通过画图就能构造出抽象的半序集。§6.7 中

图3 是正方形群的所有子群的半序集的图.

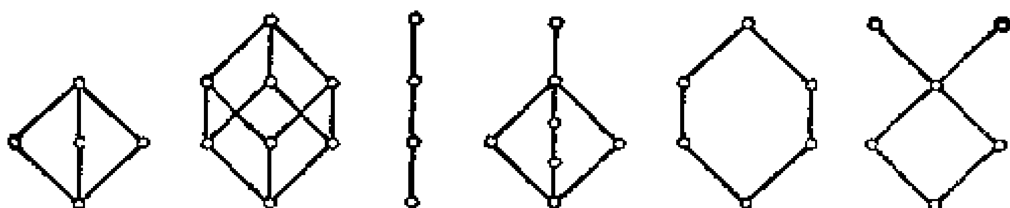


图 3

显然,在任意半序集中,关系≥也满足自反律、反对称律和传递律(这只不过是从右到左来读这些公设). 因此,由关系 $a\leqslant b$ 定义的半序集公设能够证明的任何命题,当每处的 $a\leqslant b$ 用其相反的关系 $a\geqslant b$ 来代替时,通过一系列同样的推理仍然可以证明它成立. 反之亦然. 这就是

对偶原理 *在每个半序集中成立的任何定理,如果把定理叙述中的所有符号≤都与符号≥互换,那么定理仍然成立.*

这里要强调指出,这个原理不是关于半序集的通常意义下的定理,而是关于定理的定理. 因此,它属于"元数学"的范畴.

习　　题

1. 详细说明,图 3 中的第二图怎样表示了三点集 I 的所有子集的代数.

2. 画出下列各半序集的图:

(a) 四点集的所有子集的布尔代数.

(b) 12 阶循环群的所有子群的集合.

(c) 四元数群的所有子群的集合.

(d) 整数 1, 2, 3, 4, 6, 8, 12, 24 在整除性之下构成的半序集.

(e) 54 阶循环群的所有子群的集合.

(f) 模 40 整数的环 $\mathbf{Z}_{40}$ 的所有理想的集合.

3. 证明: 习题 2 的(d), (e), (f)各部分的半序集在适当规定的意义之下都"同构".

4. 下列集合中哪些是半序集?

(a) 实数域 $\mathbf{R}$ 的所有子域,在包含关系之下.

(b) 所有数对(a, b)，如果$(a, b)\leqslant(a', b')$的意思是 $a\leqslant a'$ 和 $b\leqslant b'$.

(c) 所有实数对(a, b)，如果$(a, b)\leqslant(a', b')$ 的意思是或者 $a<a'$ 或者 $a=a'$，并且 $b\leqslant b'$.

(d) 所有实数对(a, b)，如果$(a, b)\leqslant(a', b')$的意思是 $a\leqslant a'$ 和 $b\leqslant b'$.

(e) 已知整环的所有子整环，在包含关系之下.

(f) $F[x]$ 中的所有多项式，如果 $f(x)\leqslant g(x)$ 的意思是 $f(x)$ 可整除 $g(x)$.

5. 考虑具有关系 $a<b$ 的元素系统，关系 $a<b$ 满足传递律和非自反律（$a<a$ 永远不成立）. 证明：如果 $a\leqslant b$ 的意思是或者 $a<b$ 或者 $a=b$，那么集合是半序集.

6. 证明正文中叙述的引理.

7. (a) 证明：§ 11.3 的例 1 中，格 L 是通过子空间之间的集合包含关系来定义的.

(b) 叙述并证明：关于 § 11.3 例 2 的一个类似的命题.

(c) 如果用 $m|n$ 来定义全体正整数集合的半序，那么$\wedge$和$\vee$都意味着什么呢？

§ 11.7 格

相容性原理指出怎样通过并和交来定义包含，现在我们反过来说明，可通过包含来定义并和交. 也就是说，$x\vee y$是既包含 x 又包含 y 的最小集合，而 $x\wedge y$ 是既包含在 x 中又包含在 y 中的最大集合. 这一说法是由 C. S. 皮尔斯(Peirce)提出来的，我们把它更确切地叙述如下.

设X是半序集P的某些元素的集合，如果一个元素 a，对所有 $x\in X$ 都满足 $a\leqslant x$，那么称 a 是X的“下界”. 象第四章所描述的那样，“最大下界”(g. l. b.)指的是包含其他所有下界的上界，即最大下界 c，它对其他任意下界 a，满足 $c\geqslant a$. 显然，最大下界如果存在，就一定唯一. 这是因为如果 a 和 b 都是同一个集合X的最大下界，那么 $a\geqslant b$，并且 $b\geqslant a$，因此 $a=b$.

对偶地，我们可以定义“上界”和“最小上界”(l. u. b.)，并可证明如果最小上界存在，就一定唯一．这里我们正使用了元数学的对偶原理！因此，我们可以说“集合的最大下界”、“集合的最小上界”而不用说“集合的一个最大下界”或“集合的一个最小上界”．当然这里假定这些界是存在的．

引理 1 *在任意格中，交 $x\wedge y$ 和并 $x\vee y$ 分别是由 x 和 y 两个元素组成的集合的最大下界和最小上界．*

证明 因为 $x\wedge x\wedge y=x\wedge y$ 和 $y\wedge x\wedge y=x\wedge y$，所以相容性原理指出 $x\wedge y$ 是 x 和 y 的下界．它也是最大下界，这是因为由 $z\leqslant x$ 和 $z\leqslant y$，再次根据相容性原理可推出 $z=x\wedge z=x\wedge(y\wedge z)=(x\wedge y)\wedge z$，所以 $z\leqslant x\wedge y$．因此 $x\wedge y$ 是最大下界．由对偶性就完成了整个引理的证明．

这个引理表明，任意格是具有“格性质”的半序集，所谓“格性质”是指任意两个元素具有最大下界和最小上界．我们现在将指出，这个性质完全地表征了格．

定理 3 *设 L 是任意半序集，其中任意两个元素 x，y 具有最大下界 $x\wedge y$ 和最小上界 $x\vee y$，那么在 $\wedge$ 和 $\vee$ 两种运算之下，L 是一个格，在这个格中，$a\leqslant b$ 当且仅当 $a\wedge b=a$(或等价于 $a\vee b=b$)．*

证明 只须证明幂等律、交换律、结合律和吸收律及相容性原理．而且根据对偶原理只须对最大下界来证明幂等律、交换律和结合律．由定义的对称性，交换律显然满足．因为 $x\wedge(y\wedge z)$ 和 $(x\wedge y)\wedge z$ 都是 x,y,z 三个元素的最大下界，所以满足结合律．根据定义，显然 $x\wedge x=x$，因而幂等律是显然的．为了证明相容性原理，首先假定 $x\leqslant y$，那么任意使得 $z\leqslant x$ 和 $z\leqslant y$ 的 z，满足 $z\leqslant x$；而 $x\leqslant x$ 且 $x\leqslant y$，所以 x 满足最大下界 $x\wedge y$ 的定义，因此 $x=x\wedge y$．反之，如果 $x=x\wedge y$，那么，x 是 y 的下界，所以 $x\leqslant y$，这就证明了相容性原理．吸收律可通过类似于 § 11.4 的引理 3 的证明而推

出.

上面没有提到分配律，因为分配律不是在一切格中都成立. 例如，当 x, y, z 是四元素群(图 3 的第一图)中选出的三个二阶子群时，分配律就不成立. 然而，两个与此有关的不等式成立.

定理 4 *在任意格中，半分配律成立：*

$$x\wedge(y\vee z)\geqslant(x\wedge y)\vee(x\wedge z),$$

$$x\vee(y\wedge z)\leqslant(x\vee y)\wedge(x\vee z).$$

此外，每一个分配律可推出它的对偶形式.

证明 根据对偶原理，证明可简略一半. 关于第一个半分配律，请注意，右边的两项分别是左边两项的下界；因此 x 和 $y\vee z$ 的最大下界是 $x\wedge y$ 的上界，又是 $x\wedge z$ 的上界，所以是 $x\wedge y$ 和 $x\wedge z$ 的最小上界 $(x\wedge y)\vee(x\wedge y)$ 的上界.

最后，假设 §11.3(iii) 中的第一个分配律成立，我们展开得到

$$\begin{aligned}(x\vee y)\wedge(x\vee z)&=[(x\vee y)\wedge x]\vee[(x\vee y)\wedge z]\\&=x\vee(x\wedge z)\vee(y\wedge z)\\&=x\vee(y\wedge z),\end{aligned}$$

这就是 §11.3(iii) 中的另一个分配律. 根据对偶原理，就完成了定理的证明.

由上面一些定理推得，要证明一个集合代数是布尔代数，我们只须知道，(i) 集合包含关系满足自反律、反对称律和传递律；(ii) 两个集合的并是包含这两个集合的最小集合，两个集合的交是包含在这两个集合中的最大集合；(iii) 恒等地有 $S\cap(T\cup U)=(S\cap T)\cup(S\cap U)$； (iv) 每个集合 S 有一个"补" S'，满足 $S\cap S'=O$，$S\cup S'=I$. 这也就证明了

定理 5 *布尔代数是一个分配格，这个格包含元素 O 和 I，使得对一切 a，有 $O\leqslant a\leqslant I$，并且在这个格中，每个 a 有一个补 a'，满足 $a\wedge a'=O$，$a\vee a'=I$.*

布尔代数也可以用许多其他公设系来描述．例如下面习题13就表示了一个布尔代数．

习　题

1. 图 3 中哪些图表示格？

2. 画出两个新图来表示不是格的半序集．

3. §11.6 的习题 4 中列出的集合中，哪些表示格？

4. 证明：如果在格 L 中 $b\leqslant c$，那么对一切 $a\in L$，有 $a\wedge b\leqslant a\wedge c$，和 $a\vee b\leqslant a\vee c$．

5. 叙述并证明布尔代数的对偶原理．

6. 从正文中给出的引理 1 前半部分的证明，写出后半部分的详细证明，以此为例说明对偶原理．

7. 证明：只有有限个元素的格，具有元素 O 和 I，它们满足 $O\leqslant x\leqslant I$，对一切元素 x．

*8. 证明：含有 O 和 I 的有限半序集，如果任意适合 $a_i\geqslant b_j(i, j=1, 2)$ 的元素 a_1，a_2，b_1，b_2 中间有一个元素 c，使得对所有的 i 和 j 有 $a_i\geqslant c\geqslant b_j$，那么这个半序集是一个格．

9. 链是一个全序集(即是这样的半序集，其中任意元素 x 和 y 或者满足 $x\geqslant y$，或者满足 $y\geqslant x$)．

(a) 证明：每个链是一个分配格．

(b) 证明：一个格是链当且仅当它的所有子集合都是子格．

*10. 一个格为模当且仅当由 $x\geqslant z$ 总可推出 $x\wedge(y\vee z)=(x\wedge y)\vee z$．

(a) 证明：每个分配格是模．

(b) 画出表示五元素的格而不是模的图形．

(c) 证明：下列各集合是一个模格．

(i) 一个矢量空间的所有子空间．

(ii) 一个阿贝耳群的所有子群．

(iii) 任意群的所有正规子群．

(d) 证明：在模格中，由 $x\leqslant z$ 总可推出 $x\vee(y\wedge z)=(x\vee y)\wedge z$．因此可推断，对于模格，对偶原理成立．

*11. 在任意布尔代数中，两个元素 x 与 y 的对称差定义为 $x+y=(x\wedge$

$y')\vee(x'\wedge y)$.

(a) 如果 x 和 y 是集合，它们的对称差指的是什么？画图说明。

(b) 说明 $x+y$ 满足结合律、交换律，并且具有零元素。

(c) 证明：如果把对称差当作和，把交当作积，那么每个布尔代数是一个交换环。

12. (a) 证明：如果我们在矢量空间 $\mathbf{Z}_2^n$ 中用 $\xi\eta=(x_1y_1, \cdots, x_ny_n)$ 来定义乘法，那么 $\mathbf{Z}_2^n$ 变为一个交换环，在这个环中，对一切 ξ 有 $\xi^2=\xi$.

*(b) 证明：在运算 $\xi\wedge\eta=\xi\eta$，$\xi\vee\eta=\xi+\eta+\xi\eta$，$\xi'=(1, 1, \cdots, 1)-\xi$ 之下，这个环是一个布尔代数.

*13. 证明：如果 L 是具有泛界 O 和 I 的格，在这个格中每个元素 a 都有补 a'，具有性质

$$x\leqslant a' \quad \text{当且仅当} \quad a\wedge x=O,$$
$$y\geqslant a' \quad \text{当且仅当} \quad a\vee y=I.$$

那么 L 是一个布尔代数.（提示：为了证明第一分配律，只须证明

$$e\equiv[a\wedge(b\vee c)]\wedge[(a\wedge b)\vee(a\wedge c)]'=O$$

把 e 写成交，并考虑每一项.）

§11.8 集合表示

§11.5 的主要结论是，对于布尔代数所假定的公设可推出一组恒等式，这些恒等式对于交、并、补的集合代数来说都是正确的. 实际上，已经证明了，一个特殊集合 $\mathbf{Z}_2^n$ 的适当一族子集合 $S_1, \cdots, S_n$ 具有性质：对于两个布尔多项式 p, q, $p(S_1, \cdots, S_n)=q(S_1, \cdots, S_n)$ 当且仅当它们具有相同的析取标准型. 对于给定的 n，所有这些析取标准型组成的布尔代数称为具有 n 个生成元的自由布尔代数.

我们现在将证明一个更强的结果，并顺便指出，用来定义分配格的公设完全地表征了集合的交和并的性质. 为此目的，我们需要同态和同构的概念，它们类似于对于群用过的同态和同构概念.

定义 一个从格 L 到格 M 的函数 $f: L\to M$，如果对一切 x,

$y\in L$, 有 $f(x\wedge y)=f(x)\wedge f(y)$ 和 $f(x\vee y)=f(x)\vee f(y)$, 那么函数 f 称为同态. 一一映上的同态称为同构.

例如, 由维恩图(图 1)的三个圆 X,Y,Z 生成的布尔代数与 $\mathbf{Z}_2^3$ 的所有子集组成的代数同构, 相应的函数就象 § 11.5 中所定义的那样.

引理 1 两个布尔代数(看作格)之间的同构 $f:A\longleftrightarrow B$ 一定把 A 中的泛界 O, I 和补映射到 B 中相应的泛界和补.

证明 显然, "对一切 $x\in A$, $O\wedge x=O$" 可推出 "对一切 $f(x)\in B$, $f(O)\wedge f(x)=f(O\wedge x)=f(O)$" 因此 $f(O)$ 是 B 的泛下界; $f(I)=I$ 的证明类似. 因此 "在 A 中 $x\wedge x'=O$" 可推出 "在 B 中 $f(x)\wedge f(x')=f(x\wedge x')=f(O)=O$", 对偶地有 $f(x)\vee f(x')=I$, 这就证明了 $f(x')=[f(x)]'$, 从而完成了引理 1 的证明.

定义 集环是集合 I 的这样一族子集合: 如果这个族包含任意两个子集合 S 和 T, 那么它一定包含它们的交 $S\cap T$ 和它们的并 $S\cup T$; 集域是这样的集环: 它包含 I, 包含空集 $\varnothing$, 并且如果包含任意集合 S, 那么也一定包含 S 的补 S'.

换句话说, I 的子集构成的集域恰好是 I 的所有子集构成的布尔代数 A 的一个布尔子代数; I 的子集构成的集环恰好是 A 的子格, 这时把 A 看作分配格. 我们将证明, 每个有限分配格与集环同构, 每个有限布尔代数与某(有限)集合的所有子集构成的集域同构. 这些结论有点类似于群的凯莱定理.

在证明定理 1 的这些逆命题时, 我们还需要下面的概念.

定义 格 L 的一个元素 $a>O$, 如果由 $x\vee y=a$ 可推出 $x=a$ 或 $y=a$, 则称 a 是并-不可约的; 如果 $a<I$, 并且由 $x\wedge y=a$ 可推出 $x=a$ 或 $y=a$, 则称 a 是交-不可约的; 一个元素 p, 如果 $p>O$, 并且不存在元素 x 使得 $p>x>O$, 则称 p 为原子(atom).

引理 2 在布尔代数中, 一个元素是并-不可约的当且仅当它

是一个原子.

证明 如果 p 是一个原子，那么由 $p=x\vee y$ 推出 $x=p$ 或 $x=O$；在第二种情形中 $p=O\vee y=y$，因此 p 是并-不可约的. 反过来，如果 a 不是原子也不是 O，那么对某个 x 有 $a>x>O$. 因此

$$a=a\wedge I=a\wedge(x\vee x')=(a\wedge x)\vee(a\wedge x')=x\vee(a\wedge x'),$$

这里 $x<a$. 因为 $a\wedge x'\leqslant a$，并且由 $a\wedge x'=a$ 将推出 $x=a\wedge x=a\wedge x'\wedge x=O$，所以还有 $a\wedge x'<a$，因此表明 a 是并-可约的.

现在，对任意有限格 L 的每个元素 a，设 $S(a)$ 是 L 中所有并-不可约元素 $p_k\leqslant a$ 的集合，考虑映射 $a\longmapsto S(a)$. 我们有

引理 3 在有限格 L 中，每个元素 a 满足 $a=\bigvee_{S(a)} p_k$.

证明 对于 $a=O$，可立即得出上述结论，因为 $S(O)=\varnothing$(空集)，并且 O 是空集的最小上界. 对于任意其他的 $a\in L$，我们应用数学归纳法第二原理，设 $P(n)$ 是命题：当 L 中元素 $x\leqslant a$ 的个数是 $n=n(a)$ 时引理 3 成立. 显然如果 a 是并-不可约的，则 $P(n)$ 正确. 而如果 a 是并-可约的，也不是 O，那么 $a=x\vee y$，其中 $x<a$，$y<a$，因此 $n(x)<n(a)$，$n(y)<n(a)$. 对 n 用归纳法，由此得到 x 和 y 是并-不可约元素的并：$x=\bigvee_X p_\xi$ 和 $y=\bigvee_Y q_\eta$，因此 $a=\bigvee_X p_\xi\vee\bigvee_Y q_\eta$ 是并-不可约元素的并.

引理 4 在任意有限格 L 中，映射 $a\longmapsto S(a)$ 把 L 中的交映射到集合论中的交：$S(a\wedge b)=S(a)\cap S(b)$.

证明 根据 $a\wedge b$ 的定义，$p\leqslant a\wedge b$ 当且仅当 $p\leqslant a$ 和 $p\leqslant b$.

引理 5 在有限分配格 L 中，映射 $a\longmapsto S(a)$ 把 L 中的并映射到集合论中的并：$S(a\vee b)=S(a)\cup S(b)$.

证明 一个给定的并-不可约元素 p 包含在 $a\vee b$ 中当且仅当

$$p=p\wedge(a\vee b)=(p\wedge a)\vee(p\wedge b).$$

如果 p 是并-不可约的，则上式意味着或者 $p\wedge a=p$(即 $p\leqslant a$)或者 $p\wedge b=p$(即 $p\leqslant b$)．这表明，$S(a\vee b)$包含 p 当且仅当或者 $S(a)$ 包含 p 或者 $S(b)$ 包含 p. 而反过来显然在任意格中都正确.

证毕

引理 4 和引理 5 表明，映射 $a\longmapsto S(a)$是从 L 到集环$\mathscr{R}$上的一个同态，$\mathscr{R}$是由 L 的并-不可约元素的集合 I 的所有子集构成. 而且引理 3 表明这个映射是从 L 到$\mathscr{R}$上的一一映射．这就证明了

定理 6 任意有限分配格 L 与一个集环同构.

当 L 是有限布尔代数时，引理 2 告诉我们，每个 $a\in L$是全部原子 $p\leqslant a$ 的并．还有，根据引理 4 和引理 5，对任意 $a\in L$，有

$$S(a)\cap S(a')=S(a\wedge a')=S(O)=\varnothing,$$

和

$$S(a)\cup S(a')=S(a\vee a')=S(I)=J,$$

这里 J 是 L 中所有原子(并-不可约元素)的集合．这就是$[S(a)]'=S(a')$，所以函数 $a\longmapsto S(a)$是一个同构.

我们已经证明了映射 $a\longmapsto T(a)$是从任意布尔代数 L 到 L 的原子的子集的集域$\mathscr{F}$上的一个同构．我们现在指出$\mathscr{F}$包含 L 的原子所有集合，从而证明

定理 7 任意有限布尔代数 L 与它的原子的所有集合组成的布尔代数同构.

证明的完成 这里只剩下证明下面的事实：如果 S 和 T 是 L 的原子 p_σ, p_τ，…的两个不同集合，那么 $\bigvee_S p_\sigma\neq\bigvee_T p_\tau$．但这是下面引理的推论.

引理 6 如果原子 $q\leqslant\bigvee_S p_\sigma$，那么 $q\in S$.

因为假定引理6成立，则 $\bigvee_S p_\sigma$ 只包含 S 中的原子，而不包含其他元素.

引理的证明 根据一般分配律，有

$$q=q\wedge\bigvee_S p_\sigma=\bigvee_S(q\wedge p_\sigma)$$

因为 q 是并-不可约的，所以推出上式右边有某一个 $q\wedge p_\sigma=q$，因此 $O<q\leqslant p_\sigma$，因为 p_σ 是原子，这就推出 $q=p_\sigma$.

习　题

1．证明：如果两个有限集合 I 和 J 的元素个数相同，那么 I 的所有子集组成的代数与 J 的所有子集组成的代数同构.

2．证明：对每个正整数 n，存在含有 2^n 个元素的布尔代数.

3．证明：n 元素集合的所有子集组成的布尔代数恰好有 $n!$ 个自同构.

4．(a) 求出一个从布尔代数 A 到布尔代数 B 上的格同态 $f:A\to B$，这个同态不保持泛界或补.

(b) 证明：这样的同态保持补当且仅当它保持泛界.

5．(a) 所有正整数的集合 $\mathbf{Z}^+$ 在半序"$m\leqslant n$ 当且仅当 $m|n$"之下是一个格.

(b) 证明：这个格是分配格.

(c) 鉴定它的并-不可约元素.

6．证明：如果有限分配格 L 的全体并-不可约元素是链 C，那么 L 本身是一个链. L 的元素比 C 的元素多多少?

第十二章　超限算术

§12.1　数与集合

本章将讨论数与集合之间的联系．这就是对正整数用基数的方法处理，它与§2.6中用皮亚诺公设阐明的序的处理方法对比，后者是把大家所熟悉的序列"一，二，三，四，…"看作是基本的．这种基数处理方法能够使我们按照集合来定义数，从而减少在数学中必须假定的不加定义的术语的总数．但是为了实现这一方案，还需要很大变动的基本概念来与本书吻合．

所以，我们将假定读者已熟悉正整数和集合的概念，并从这里进行讨论．我们的目的是，推广这种基数方法以便给出无限基数的严格定义，无限基数概念在现代数学中起着很基本的作用．用这个定义，我们指出基数如何相加、相乘、并产生任意基数幂，在这过程中证明，这些运算具有正整数相应运算的绝大部分（虽然不是全部）性质．

数与集合之间的关系来源于下面定义．

定义　*设 n 是任意正整数．一个集合 S 称为具有基数 n（用符号表示就是 $o(S)=n$①）当且仅当 S 的全体元素与整数 $1,2,3,\cdots,n$ 之间存在双射．*

这个定义意味着 S 的全体元素可以标明 $s_1, s_2, s_3, \cdots, s_n$，其中 s_k 是 S 中对应于整数 k 的元素．换句话说，我们可以来数 S 的元素，一直数到 n，每个元素数一次而且仅数一次．由此得到一个

① 有时称空集为具有零基数的集．

推论：如果两个集合 S 和 T 具有相同的基数，那么 S 和 T 之间存在一个双射，即存在对应关系 $s_1 \longleftrightarrow t_1, \cdots, s_n \longleftrightarrow t_n$. 但是下面事实并不显然：同一个集合不能有两个不同的基数，也就是说，按不同顺序重新计数，我们不能得到不同的元素总数. 我们现在就来证明这个事实，先叙述一个稍微一般的结果.

定理 1 设 m 和 n 是正整数. 集合 $\{1, 2, \cdots, m\}$ 和集合 $\{1, 2, \cdots, n\}$ 的一个真子集之间存在一个双射当且仅当 $m<n$.

证明 如果 $m<n$，那么双射：$1 \longleftrightarrow 1, 2 \longleftrightarrow 2, \cdots, m \longleftrightarrow m$ 就是我们所要的对应. 定理 1 的这一半命题比较显然，但是分析逆命题时必须更谨慎些.

当 $m=1$ 时，逆命题是显然的，这因为 1 是最小正整数；因此我们可以对 m 用归纳法. 我们现在假设集合 $\{1, \cdots, m\}$ 和整数集合 $\{1, \cdots, n\}$ 的真子集 S 之间有一个双射 $1 \longleftrightarrow f(1), \cdots, m \longleftrightarrow f(m)$. 定义一个新的双射 $i \longleftrightarrow g(i)$ $(i=1, \cdots, m-1)$ 如下：

$$\begin{aligned} g(i) &= f(i), \qquad \text{当 } f(i) \neq n \\ g(i) &= f(m), \qquad \text{当 } f(i) = n \end{aligned} \tag{1}$$

因为至多对一个 i 使 $f(i)=n$，所以对应 $i \longleftrightarrow g(i)$ 将是整数 $1, \cdots, m-1$ 和整数 $1, \cdots, n-1$ 中的某些整数之间的一一对应.

根据假定，所有整数 $f(i)$ 的集合 S 是集合 $\{1, \cdots, n\}$ 的真子集，这意味着 S 不包含所有整数 $1, \cdots, n$. 让我们在 S 之外选取一个最小的正整数 $k \leqslant n$，所以对 $i=1, \cdots, m$，$f(i)$ 决不等于 k. 当 $k<n$ 时，条件(1)表明，没有一个 $g(i)$ 等于 k；当 $k=n$ 时，$f(i)=n$ 决不成立，所以没有一个 $g(i)$ 等于 $f(m)$. 无论哪种情况，整数 $g(1), \cdots, g(m-1)$ 不能包含所有整数 $1, \cdots, n-1$，所以 $i \longleftrightarrow g(i)$ 是整数集合 $\{1, \cdots, m-1\}$ 和整数集合 $\{1, \cdots, n-1\}$ 的真子集之间的一一对应. 现在根据数学归纳法假设，我们能得到 $m-1<n-1$，因此两边都加上 1，就有 $m<n$.

推论 1 集合$\{1, \cdots, m\}$和集合$\{1, \cdots, n\}$的一个子集合之间存在一个双射当且仅当$m \leqslant n$.

证明 如果$m \leqslant n$, 那么双射$1 \longleftrightarrow 1, \cdots, m \longleftrightarrow m$就是所要的对应. 反过来, 如果$i \longleftrightarrow f(i)$是集合$\{1, \cdots, m\}$和整数$1, \cdots, n$中某些整数之间的双射, 那么它是集合$\{1, \cdots, m\}$和$\{1, \cdots, n, n+1\}$的真子集之间的双射. 因此根据定理 1, 有$m < n+1$, 所以$m \leqslant n$.

推论 2 如果集合$\{1, \cdots, m\}$和集合$\{1, \cdots, n\}$之间存在一个一对一映上, 那么$m = n$.

这因为, 根据推论 1, 有$m \leqslant n$和$n \leqslant m$, 因此$m = n$. 这就表明, 同一个集合不能有两个不同的正整数作为它的基数.

推论 3 如果S是集合$\{1, \cdots, n\}$的一个真子集, 那么在集合$\{1, \cdots, n\}$和集合S之间不存在双射.

证明 如果存在这样的双射, 由定理 1 将得出$n < n$, 这是矛盾的.

上述结果可直接推出下面结论. 设S和T是任意两个集合, 它们的基数分别为正整数m和n. 那么$m \leqslant n$当且仅当S和T的子集合之间存在一个双射; $m = n$当且仅当S和整个T之间存在一个双射.

习　题

1. 证明: 如果集合S的基数为n, t是S中的特定元素, 那么S和$\{1, \cdots, n\}$之间存在一个双射, 使得t对应于n.

2. 证明: 如果集合S的基数为n, 那么从S中去掉一个元素后, 便留下一个基数为$n-1$的集合S^*.

3. 用证明定理 1 时用过的方法直接证明推论 1.

4. 象习题 3 那样证明推论 3.

§12.2 可 数 集

一个集合称为有限集当且仅当它的元素可以用通常的方法计数. 下面我们把这个概念阐述得更确切些.

定义 一个非空集合S称为有限集当且仅当它的基数是一个正整数. 不是空集也不是有限集的集合称为无限集.

例如, 所有正整数的集合$\mathbf{Z}^+$是无限集. (利用定理1, 这是不难证明的.)我们现在引进如下概念: 无限集也可以看作是有基数的.

定义 一个集合S, 如果它与所有正整数的集合是双射的, 那么称S为可数集, 或称S具有基数$\mathbf{d}$(用符号表示①就是$o(S)=\mathbf{d}$).

这个定义的条件等价于: 可以把S的全部元素列成普通的无限序列$s_1, s_2, s_3, \cdots, s_n, \cdots$, 使得$S$的每个元素出现一次且仅出现一次. 如果另一个集合$T$与可数集$S$是双射的, 那么可推出$T$本身也是可数的.

定理2 (伽里略(Galileo)悖论) 任意可数集都有一个可以把它映射到它的真子集上的双射.

证明 可数集(比如说集合S)的所有元素可以根据定义写成序列$s_1, s_2, s_3, \cdots$, 它以全体不同正整数作为下标. 双射$s_1\longleftrightarrow s_2$, $s_2\longleftrightarrow s_3$, $\cdots$, $s_i\longleftrightarrow s_{i+1}$, $\cdots$是集合S和从S中删去s_1而得到的集合之间的一一对应. 证毕

可以看出$\mathbf{d}$("可数无穷")是最小的无限基数. 更确切地说, 这是

定理3 任意无限集包含一个可数子集.

证明 设S是无限集, 在S中选取任意元素s_1. 然后从

① 常常用希伯来字母$\aleph_0$(读作"阿勒夫-零")代替符号$\mathbf{d}$.

$S-\{s_1\}$中选取第二个元素 s_2；再从 $S-\{s_1, s_2\}$中选取第三个元素 s_3，等等．因为 S 是无限的，所以 $S-\{s_1, s_2, \cdots, s_n\}$决不可能是空集，因此我们总可以在这里选取一个元素①s_{n+1}；并且这个过程永远不会停止，直到我们构造出 S 的不同元素的无限序列．

推论（戴德金-皮尔斯）一个集合 S 是无限集当且仅当有一个把 S 映射到它的真子集上的双射．

证明 如果 S 是基数为 n 的有限集，那么 S 与$\{1, \cdots, n\}$是双射的，所以定理 1 的推论 3 断言 S 不能与它的一个真子集双射．相反地，设 S 是任意无限集．它将包含由元素 $u_1, u_2, u_3, \cdots$组成的可数子集U．构造一个函数，它把 U 中每个元素 u_i 与其后继 u_{i+1} 对应起来，再把 S 中不属于 U 的每个元素与其自身对应起来，这个函数就是从 S 到 S 的真子集的双射．　　证毕

实际上，很多无限集原来都是可数集（具有基数 $\mathbf{d}$）．下面定理给出两个例子．

定理 4 所有整数集合 $\mathbf{Z}$ 是可数集；所有有理数的集合 $\mathbf{Q}$ 是可数集．

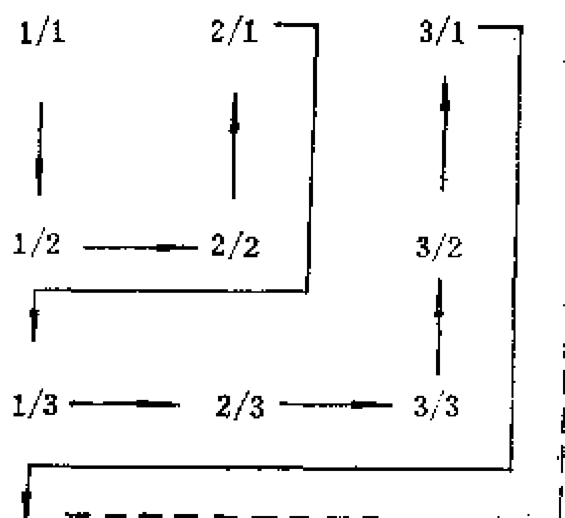


图 1

证明 对应$n\longleftrightarrow 2n+1(n=0, 1, 2, \cdots)$，$(-n)\longleftrightarrow 2n(n=1, 2, 3, \cdots)$是所有整数的集合$\{0, -1, 1, -2, 2, \cdots\}$和所有正整数集合$\{1, 2, 3, 4, 5, \cdots\}$之间的一一对应．这就证明了第一个断言．

下面我们来证明，所有正有理数的集合 $\mathbf{Q}^+$是可数集．为此我们首先把所有正整数的商排成一个无限正方形，如图 1 所示．如

① 这个构造用了集合论中称为选择公理的一个基本原理：给定任意集合 S，存在一个"选择函数"γ，它从任意非空集合 $T\subset S$ 中选取一个元素 $\gamma(T)\in T$．

果按照顺序，围绕较小的正方形边界来排列，那么我们可以把所有这样的商排成下面普通的无限序列. 第一项是$\frac{1}{1}$; $\frac{n}{1}$的后继是$\frac{1}{n+1}$; $\frac{m}{n}$的后继，当 $m<n$ 时，是$\frac{m+1}{n}$; 当$m\geqslant n>1$时，是$\frac{m}{n-1}$. 从这个序列中删去所有不是既约分数的分数(或者等价地说，删去的这些分数等于前面已列举过的另一些整数的商). 所得到的子序列把全体正有理数列成一个普通的序列，并建立起$\mathbf{Q}^+$和$\mathbf{Z}^+$之间的双射$\frac{m}{n}\longleftrightarrow k$. 而这又可以很容易地扩张成所有有理数的集合$\mathbf{Q}$和所有整数的集合$\mathbf{Z}$之间的双射$\frac{m}{n}\longleftrightarrow k$, $0\longleftrightarrow 0$, $-\frac{m}{n}\longleftrightarrow -k$. 因为$\mathbf{Z}$是可数集，所以推出$\mathbf{Q}$是可数集.

习　题

1. 证明：7的所有整数倍数组成的集合是可数集.

2. 证明：有理数域上的有限维空间 $\mathbf{Q}^n$ 中所有矢量组成的集合是可数集.

3. 直接证明：所有正整数的集合与一个有限集之间的双射是不存在的.

4. 证明：如果T和U是可数集，那么$S=T\cup U$是可数集.

5. 证明：如果$S=T\cup U$，其中S是可数集，T是有限集，那么U是可数集.

6. 证明：可数集的每个子集或者是有限集或者是可数集.

7. 证明：仅由数字 9 组成的所有十进小数构成一个无限序列，这样的十进小数有可数多个.

8. 分别建立所有整数的集合和它的三个真子集之间的特殊的双射.

9. 证明：在图 1 中，当 $m\leqslant n$ 时，$\frac{m}{n}$是第$(n-1)^2+m$项；当 $m>n$ 时，$\frac{m}{n}$是第 m^2-n+1 项.

10. 证明：域 $\mathbf{Q}(\sqrt{2})$ 是可数的(参看 § 2.1).

11. 证明：每个群包含一个可数子群或有限子群.

12. 列出实数域和它的真子集之间的双射.

13. 证明：所有形为 $r+r'\sqrt{-1}$ $(r, r'\in\mathbf{Q})$ 的数组成的集合是可数的.

*14. 证明：所有有理系数多项式组成的环 $\mathbf{Q}[x]$ 是可数的.

§12.3 其他基数

不是所有的无限集合都是可数的：存在不止一个的“无限”基数. 例如

定理 5 （康托(Cantor)）所有实数的集合 $\mathbf{R}$ 不是可数的.

证明 我们应用所谓“对角线法”. 假定所有实数有一个排列 $x_1, x_2, x_3, \cdots$. 把这些数的小数点后的十进小数展开式按它们的排列顺序排成一个正方形阵列，如图 2 所示. 由这个阵列的对角线

$$x_1=\cdots.\ a_{11}a_{12}a_{13}a_{14}\cdots$$
$$x_2=\cdots.\ a_{21}a_{22}a_{23}a_{24}\cdots$$
$$x_3=\cdots.\ a_{31}a_{32}a_{33}a_{34}\cdots$$
$$x_4=\cdots.\ a_{41}a_{42}a_{43}a_{44}\cdots$$

图 2

上的数字按以下方式构造一个新的数 b（它在 0 和 1 之间）：这里 a_{nn} 是对角线上的第 n 个数字，设 b_n 是 b 的小数展开式的第 n 个数字，取

$$b_n=\begin{cases}a_{nn}-1, & \text{当 } a_{nn}\neq 0,\\ 1, & \text{当 } a_{nn}=0,\end{cases}$$

那么 $b=0.b_1b_2b_3b_4\cdots$是某个实数的十进小数 b 的展开式，这个实数不同于上述排列中的第 n 个数 x_n，因为至少小数点后第 n 位数字不相等. 于是没有一个 x_n 等于 b，这与我们的假定“上述排列包含所有实数”相矛盾.

注记 对于下述情形，这个证明就复杂了：有一些数，例如 $1.000\cdots=0.999\cdots$，可以有两种不同的小数展开式，一种是以无穷个连续的 9 为结尾，另一种是以无穷个连续的 0 为结尾. 当我们

假定原来排列中的小数 $x_1, x_2, \cdots$ 都不用第一类展开式(带 9 的)，这个麻烦就可以避免. b 的构造永远不会产生数字 $b_n=9$, 因此 b 的小数展开式是与 x_n 的小数展开式进行比较的合适的形式.

定义 与所有实数的集合 $\mathbf{R}$ 双射的集合 S 称为具有连续统的基数 $\mathbf{c}$(用符号表示就是 $o(S)=\mathbf{c}$).

实际上，几何学和数学分析中出现的大部分集合具有基数 $\mathbf{d}$ 和 $\mathbf{c}$. 可以用几个特殊的结构分别说明这一点，但从长远来看还是先证明施罗德 (E. Schroeder) 和伯恩斯坦 (F. Bernstein) 的一般原理更容易些. 阐述这个原理时，包含着基数的一般概念，我们现在就来定义它.

定义 集合 S 的基数是所有可以双射到 S 的集合 (set) 组成的类①(class); S 的基数用 $o(S)$ 表示.

由此得出，两个集合 S 和 T 具有相同基数(或基数等价)当且仅当 S 和 T 之间有双射. 我们用符号等式 $o(S)=o(T)$ 来表示这个结论.

由于 §12.1 最后一句话，基数之间的不等式概念可以与普通正整数之间不等式的概念一致.

定义 当集合 S 和集合 T 之间有一个由 S 到 T 的单射时，我们就称集合 T 基数优于集合 S, 并且记作 $o(S)\leqslant o(T)$.

定理 6 (施罗德-伯恩斯坦)如果 $o(S)\leqslant o(T)$, 且 $o(T)\leqslant o(S)$, 那么 $o(S)=o(T)$.

换句话说就是，如果存在一个由 S 到 T 的单射，还存在另一个由 T 到 S 的单射，那么就存在整个 S 和整个 T 之间的双射(逆命题是显然的).

证明 设 $s\longmapsto s\tau$ 是已知的由 S 到 T 的单射，且设 $t\longmapsto t\sigma$

---

① 这个概念好象"化学元素"的概念，同样是抽象概念，化学元素指的是具有一个确定核电荷(即具有一个确定结构)的所有原子.

是已知的由T到S的子集的单射. S的每个元素s至多是T的一个元素$t=s\sigma^{-1}$的象$t\sigma$; 这个元素t(如果存在的话)本身在S中也至多有一个象源$s'=t\tau^{-1}=s\sigma^{-1}\tau^{-1}$, 等等. 用这个方法尽可能地追溯$S$的每个元素的"祖先"(并且对$T$的每个元素也这样做), 我们看出有三种可能情况: (a)类元素, 它的"祖先"可以无止境地追溯下去, 也许是周期地追溯下去(见习题 13); (b)类元素,是S中的"无母祖先"传下来的; (c)类元素, 是由它在T中的"无母祖先"传下来的. 对应着这三种情况, 我们把S分成子集合S_a, S_b, S_c, 把T分成子集合T_a, T_b, T_c. 而且, 包含S或T的任意元素的类一定包含它的"祖先"和"后裔".

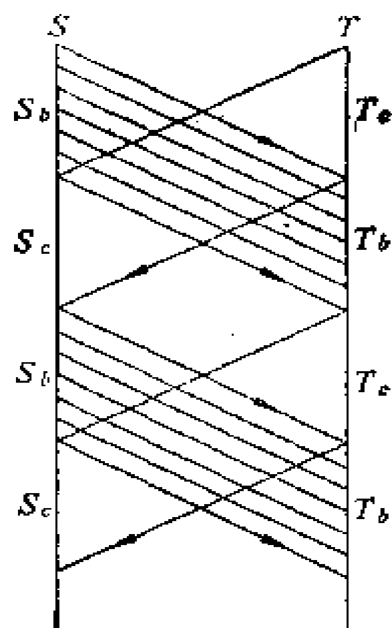


图 3

实际上, σ显然是(τ也是!)S_a和T_a之间的双射, 这因为, S_a的每个元素是T_a的一个元素且只一个元素在σ之下的象, 而T_a的每个元素t是S_a的一个元素且只一个元素$t\sigma$的象源. 类似地, τ是(但σ不是!)S_b和T_b之间的双射, 而σ是(但τ不是!)S_c和T_c之间的双射. 把这三个双射合在一起, $S_a\longleftrightarrow T_a, S_b\longleftrightarrow T_b, S_c\longleftrightarrow T_c$, 我们就得到整个$S$和整个$T$之间的双射. 证毕

这些情况可用图 3 来解释, 但这个图没有标出"祖先"可以无止境追溯下去的那些元素. 集合S和T用两条垂直直线上的点来表示, 这里τ用朝右斜下方的箭头来表示, σ用朝左斜下方的箭头来表示. 而S_b到T_b的双射用不带箭头的直线来表示.

定理 7 直线段S_1: $0<x<1$, 平面上的单位正方形S_2: $0<x, y<1$, 空间中的单位立方体S_3: $0<x, y, z<1$, 都具有基数 c.

证明 函数$x \longmapsto e^x = y$(其逆为$y \longmapsto \log_e y$)是$-\infty < x < +\infty$和$0 < y < +\infty$之间的一一对应；函数$y \longmapsto \frac{y}{1+y} = z$(其逆为$z \longmapsto \frac{z}{1-z}$)是$0 < y < +\infty$和$0 < z < 1$之间的一一对应. 因此函数$x \longmapsto \frac{e^x}{1+e^x} = z$是由$-\infty < x < +\infty$到$0 < z < 1$的双射. 这就证明了第一个结论.

为了证明第二个结论，我们考虑映射

$$(0.x_1x_2x_3\cdots, 0.y_1y_2y_3\cdots) \longmapsto 0.x_1y_1x_2y_2x_3y_3\cdots, \tag{2}$$

这是写成小数形式的0,1间的实数有序对和0,1间单个实数之间的映射. 它是正方形S_2和直线段S_1的一个子集合之间的单射(虽然不连续)——如果不排除某一位后仅由数字9组成的小数，那么这个映射就是S_2和整个S_1之间的双射了. 这就证明了$o(S_2) \leqslant o(S_1)$. 但是又有$o(S_1) \leqslant o(S_2)$$\left(\text{通过明显的映射} x \longmapsto \left(x, \frac{1}{2}\right)\right)$；因此根据定理6得到$o(S_2) = o(S_1)$，这就是说$S_2$的基数是$\mathbf{c}$. 类似的映射

$$(0.x_1x_2x_3\cdots, 0.y_1y_2y_3\cdots, 0.z_1z_2z_3\cdots) \longmapsto (0.x_1y_1z_1x_2y_2z_2\cdots) \tag{3}$$

可证明$o(S_3) \leqslant o(S_1)$，因此类似地有$o(S_3) = \mathbf{c}$. 证毕

基数为$\mathbf{c}$的集合的另一些例子将在习题中给出.

习　题

1. 为什么施罗德-伯恩斯坦定理在集合T_a是空集的情况下是显然的？在这种情况中S_e怎么样？

2. 当S是区间$-1 \leqslant s \leqslant \frac{1}{2}$，$T$是区间$-1 \leqslant t \leqslant \frac{1}{2}$，$\tau$是单射$s \mapsto s^3$，$\sigma$是双射$t \mapsto t^3$时，明显地确定集合$S_a$，$S_b$，$S_c$，$T_a$，$T_b$，$T_c$.

3. 当S是正整数集合，T是非负整数集合，τ是映射$s \mapsto s$，σ是映射$t \mapsto t+1$，确定集合S_a，S_b，S_c，T_a，T_b，T_c.

4. 证明：n 维实空间中任意包含连续弧的子集合都具有基数 $\mathbf{c}$.

5. 证明：如果存在一个从集合 S 到整个第二个集合 T 的满射，那么 $o(T)\leqslant o(S)$.

6. 证明习题 5 的逆命题：如果 $o(T)\leqslant o(S)$，那么存在一个 S 到 T 上的满射.（你可以假定选择公理成立.）

7. 基数等价关系（$o(S)=o(T)$）满足自反律、对称律和传递律吗？关系 $o(S)\leqslant o(T)$ 满足这些定律吗？给出证明.

8. 证明：如果 $o(S)\leqslant o(T)$ 且 $o(U)=o(S)$，那么 $o(U)\leqslant o(T)$.

*9. 证明：以四元数为元素的 $n\times n$ 矩阵共有 $\mathbf{c}$ 个.

*10. 在 0 和 1 之间所有实数的集合与所有无限小数 $0.a_1a_2a_3\cdots$ 的集合之间建立一个明显的双射.

*11. 建立区间 $0<x<1$ 和 $0\leqslant x\leqslant 10$ 之间的双射.

*12. 不用施罗德-伯恩斯坦定理直接证明：

(a) 所有非负实数的集合具有基数 $\mathbf{c}$.

(b) 那些非整数的正实数组成的集合具有基数 $\mathbf{c}$.

*13. 设 N 是自然数集合，当 $S=N\cup\{u,v,w\}$，$T=N\cup\{u,v,w\}$，$\sigma(n)=n+1$，并且在 $\{u,v,w\}$ 上循环（即 $\sigma(u)=v$，$\sigma(v)=w$，$\sigma(w)=u$）；$\tau(n)=n+2$，并且在 $\{u,v,w\}$ 上恒等（即 $\tau(u)=u$，$\tau(v)=v$，$\tau(w)=w$），明显地确定集合 S_a，S_b，S_c，T_a，T_b，T_c.

*§12.4 基数的加法与乘法

无限基数就象有限基数那样可以相加和相乘，除了消去律以外的所有定律仍然成立.

设 m 和 n 是正整数，我们可以用下面方法构造一个基数为 $m+n$ 的集合：基数为 m 的集合 S'（比如说集合 $\{1,2,\cdots,m\}$）再加上一个基数为 n 的与它不相交的集合 S''（比如说集合 $\{m+1,m+2,\cdots,m+n\}$）. 那么并 $S'\cup S''$ 就具有基数 $m+n$. 类似地，所有数对 (i,j) 的集合（其中 i 跑遍整数 $1,\cdots,m$，j 跑遍整数 $1,\cdots,n$；例如，$m\times n$ 矩阵的元素的全体下标）具有基数 mn. 我们并不去证明这些熟悉的事实，而要指出这些事实暗示着基数加法和乘法

运算可以推广到无限基数，如下所述.

定义 设α和β是任意基数. 则分别具有α个和β个元素的两个不相交子集的并的基数是$\alpha+\beta$，所有数对(x, y)组成的集合（这里，x跑遍具有α个元素的集合，y跑遍具有β个元素的集合）的基数是$\alpha\beta$.

基数加法是单值的，这是因为如果S是两个不相交子集S'和S''的并，T是两个不相交子集T'和T''的并，并且S'和T'之间，S''和T''之间都存在双射，那么我们可以把这两个双射合并成整个S和整个T之间的一个双射. 类似地，基数的乘法也是单值的. 事实上，普通算术的大部分定律，如同用于有限数的情形一样，都可用于无限数的情形①.

定理 8 基数加法和乘法满足交换律和结合律；乘法对于加法满足分配律；1是单位元素.

证明 基数加法的交换律和结合律是布尔代数定律的推论. 因为不管对什么集合S和T，函数$(x, y)\longmapsto(y, x)$是所有数对$(x, y)[x\in S, y\in T]$的集合和所有数对$(y, x)[y\in T, x\in S]$的集合之间的双射，由此推出乘法交换律. 显然，所有3-数组$((x, y), z)$ $[x\in S, y\in T, z\in U]$的集合到所有3-数组$(x, (y, z))[x\in S, y\in T, z\in U]$的集合存在一个双射，这里$S, T, U$是任意集合，由此推出乘法结合律. 最后，如果$T$和$U$是不相交的，那么所有数对$(x, w)$ $[x\in S, w\in T$或$U]$的集合的基数显然是$o(S)[o(T)+o(U)]$；而所有数对$(x, y)[x\in S, y\in T]$的集合与所有数对$(x, z)[x\in S, z\in U]$的集合的并的基数是$o(S)o(T)+o(S)o(U)$. 这两个集合之间存在一个明显的双射，因此这就证明了分配律. 对任意基数α，显然有

---

① 可惜，根据以下定理（我们不证明），这个事实就不那么重要了，这个定理是说，任意两个无限基数的和或积只是这两个基数中较大的一个. 而超限取幂（§12.5）更为重要.

$1\cdot\alpha=\alpha$.

定理 9 *加法消去律和乘法消去律对无限基数来说是不成立的.*

证明 定理 2 的证明表明 $\mathbf{d}=\mathbf{d}+1$. 但由此可推出 $\mathbf{d}+1=(\mathbf{d}+1)+1=\mathbf{d}+2$, 虽然 $1\neq 2$, 这就说明加法消去律不成立. 再有, 正整数的集合 $\mathbf{Z}^+$ 可以分成不相交的偶数集合和奇数集合, 它们都是可数的, 因此 $\mathbf{d}+\mathbf{d}=\mathbf{d}$. 所以, 根据定理 8, 有 $(1+1)\mathbf{d}=1\cdot\mathbf{d}$ 或 $2\mathbf{d}=1\mathbf{d}$, 还是 $2\neq 1$. 证毕

实际上, 对于所有无限基数, 方程 $\alpha=\alpha+1$ 和 $\alpha=\alpha+\alpha$ 成立, 但我们不去证明它们.

由此得出一个推论, 有限基数和无限基数的系统不能被嵌入任何一个具有减法和除法的系统中去. 你能证明它吗?

习　题

1. 详细证明(用布尔代数): 基数加法满足交换律和结合律.
2. 证明: 对任意无限基数 α, 有 $\alpha=\alpha+1$. (提示: 用定理 3.)
3. 证明: $\mathbf{d}+\mathbf{d}+\mathbf{d}=\mathbf{ddd}=\mathbf{d}$. (提示: 见图 1.)
4. (a) 证明: 如果 n 是有限基数, 那么 $\mathbf{d}+n=\mathbf{d}$.

   (b) 同样证明: $n\mathbf{d}=\mathbf{d}$.
5. 不用定理 6 证明: $\mathbf{c}+\mathbf{d}=\mathbf{c}$.
6. 不用 §12.5 证明: $\mathbf{c}+\mathbf{c}=\mathbf{cc}=\mathbf{c}$.
7. 证明: $\mathbf{dc}=\mathbf{c}$.
8. 证明 §12.4 中最后一个命题.
9. (a) 证明: 如果 $x\geqslant\mathbf{d}$, 那么 $x+\mathbf{d}=x$.

   (b) 证明: 如果 $x+\mathbf{d}=\mathbf{c}$, 那么 $x=\mathbf{c}$.

*10. 对于可数群 G, 考虑关于 G 的可能有限子群 S 的阶的拉格朗日定理(第六章)的证明.

(a) 证明: 证明中对 S 的阶不加限制.

(b) 证明: 在可数群 G 中可存在任意给定有限阶的子群.

*§12.5 取　　幂

如果S和T是有限集合，它们的基数分别为$m=o(S)$和$n=o(T)$，那么普通幂$n^m=o(T)^{o(S)}$可以描述为由集合S到T的函数的个数．对任意这样的对应$x\longmapsto y$确定一个函数$y=f(x)$，它对每个自变量$x\in S$，都赋给一个值$y\in T$．为了计算所有不同的抽象函数f的个数，我们注意，S的第一个元素x恰好有$o(T)$种可能的象；对于每一个这样的象，S的第二个元素的象y有$o(T)$种选择，等等．所以所有$o(S)$个元素的象的选法总数是$o(T)$自乘$o(S)$次，即等于$o(T)^{o(S)}$．

$o(T)^{o(S)}$的这种组合特征可以应用于无限基数．

定义　设α和β是任意非零基数．则β^α是从由α个元素组成的集合到由β个元素组成的集合的函数的个数．

这定义了一个单叶运算，即如果$\alpha=\alpha'$且$\beta=\beta'$，那么$\beta^\alpha=\beta'^{\alpha'}$，其证明是显然的，我们把它略去．

定理 10　$\mathbf{c}=2^{\mathbf{d}}$

证明　0和1之间的每个实数x有一个二进数展开式$0.x_1x_2x_3\cdots$，它可看作一个无穷序列$x_1,x_2,x_3,\cdots$，其中x_i等于0或1．不同的实数x和y具有不同的展开式（§4.3），因此函数$f(x)=(x_1,x_2,x_3,\cdots)$是一一的．可是这种序列的个数由定义是下面函数的个数，这种函数是从可数的定义域（即，序列中所有$\mathbf{d}$个位置组成的集合）到两个元素（即0和1）组成的取值域上的函数．我们推出，0和1之间的实数至多有$2^{\mathbf{d}}$个，因此根据定理7有$\mathbf{c}\leqslant 2^{\mathbf{d}}$．

另一方面，每个仅由数字3和7组成的无限十进小数也可表示不同的实数，因此$2^{\mathbf{d}}\leqslant\mathbf{c}$．现在利用定理6，我们就得到$\mathbf{c}=2^{\mathbf{d}}$．

定理 11　对于任意基数α,β和γ，下面的取幂法则成立：

(i) $\alpha^\beta\alpha^\gamma=\alpha^{\beta+\gamma}$，　　　(ii) $(\alpha\beta)^\gamma=\alpha^\gamma\beta^\gamma$，

(iii) $(\alpha^\beta)^\gamma=\alpha^{\beta\gamma}$, (iv) $\alpha^1=\alpha$ 和 $1^\alpha=1$.

证明 (iv)中的两个恒等式的证明是显然的. 为了证明恒等式(i)～(iii), 我们假定 S, T 和 U 分别是具有 α, β 和 γ 个元素的集合, 其中 T 和 U 是不相交的.

(i)的证明 设 V 是这样的集合, 在 V 中, T 和 U 是互补子集, 我们考虑由集合 V 到集合 S 的函数 $h(v)$. 根据定义, 这样函数的个数等于 $\alpha^{\beta+\gamma}$. 另一方面, 每个这样的函数确定一对无关的函数 $(f(t), g(u))$, 并且每对无关函数 $(f(t), g(u))$ 确定一个这样的函数, 这里 $f(t)$ 是由 T 到 S, $g(u)$ 是由 U 到 S. 由定义, 这种函数对的个数是 $\alpha^\beta\alpha^\gamma$.

(ii)的证明 考虑函数 $h(u)$, 它对每个 $u\in U$ 赋给一对值 $(s, t)=(f(u), g(u))$, s 和 t 分别在 S 和 T 中取任意值. 由定义可知, 这样函数的个数是 $(\alpha\beta)^\gamma$. 但是它也是函数对 $(f(u), g(u))$ 的个数 $\alpha^\gamma\beta^\gamma$, 其中 $f(u)$ 是从 U 到 S, $g(u)$ 是从 U 到 T.

(iii)的证明 考虑两个变量 $t\in T$ 和 $u\in U$ 的函数 $f(t, u)$, 它在 S 中取值, 根据定义, 这种函数的个数是 $\alpha^{\beta\gamma}$. 但是, 对每个固定的 u, $f(t, u)$ 联系一个对应关系 $f_u(t)$, 它赋给每个 t 一个值 $f_u(t)=f(t, u)\in S$. 反过来, 每个映射 $u\longmapsto f_u$ 定义了两个变量 t 和 u 的函数 $f(t, u)=f_u(t)$. 因为根据定义 f_u 的个数是 α^β, 所以 $f(t, u)$ 的个数是 $(\alpha^\beta)^\gamma$. 证毕

由定理 10 和定理 11, 我们可以从相应的关于 $\mathbf{d}$ 的方程推出一些包含 $\mathbf{c}$ 的方程. 例如,

$$\mathbf{c}^2=(2^\mathbf{d})^2=2^{2\mathbf{d}}=2^\mathbf{d}=\mathbf{c},$$

$$2\mathbf{c}=2^1 2^\mathbf{d}=2^{1+\mathbf{d}}=2^\mathbf{d}=\mathbf{c},$$

$$\mathbf{c}^\mathbf{d}=(2^\mathbf{d})^\mathbf{d}=2^{\mathbf{d}^2}=2^\mathbf{d}=\mathbf{c} \qquad \text{(参看定理 4)}.$$

利用这些结果和下面的习题 1 以及定理 6, 我们容易地得到象 $\mathbf{d}^\mathbf{d}=\mathbf{c}$ 和 $n^\mathbf{d}=\mathbf{c}$(对任意自然数 $n>1$)等这样的法则.

定理 12 对任意基数 α, 有 $\alpha<2^{\alpha}$.

说明 这个记号是指 $\alpha\leqslant 2^{\alpha}$, 但 $\alpha\neq 2^{\alpha}$.

证明 设 S 是基数为 α 的任意集合. 则 2^{α} 就是函数 $f(x)$, $g(x)$, …的个数, 这些函数的定义域是 S, 取值为 0 和 1. 定义 $f_x(y)=0$, 当 $x\neq y$; $f_x(x)=1$, 于是我们就得到 S 和从 S 到集合 $\{0,1\}$ 的函数的特殊集合之间的双射 $x\longleftrightarrow f_x$. 这就证明了 $\alpha\leqslant 2^{\alpha}$.

反过来, 设在 S 和以 S 为定义域、取值为 0 和 1 的函数之间, 给定任意双射 $x\longleftrightarrow g_x$. 构造一个新函数 $h(x)$: $h(x)=0$, 当 $g_x(x)=1$; $h(x)=1$, 当 $g_x(x)=0$. 这就定义了一个以 S 为定义域、取值为 0 和 1 的函数, 而且由构造可知, 对所有函数 g_x, $h(x)\neq g_x(x)$. 我们得出结论: h 与每个 g_x 都不同, 因此在 S 和以 S 为定义域、取值为 0 和 1 的所有函数的集合之间不存在任何双射. 用符号表示就是, $\alpha\neq 2^{\alpha}$.

习 题

1. 证明: 如果 $\alpha\leqslant\beta$, 那么对所有 γ, 有

   (a) $\alpha+\gamma\leqslant\beta+\gamma$ (b) $\alpha\gamma\leqslant\beta\gamma$

   (c) $\alpha^{\gamma}\leqslant\beta^{\gamma}$ (d) $\gamma^{\alpha}\leqslant\gamma^{\beta}$

2. 证明: $\mathbf{c}^{\mathbf{c}}=2^{\mathbf{c}}$. (提示: 用 § 12.4 习题 7.)

3. 证明: 如果集合 S 的基数为 α, 那么 S 的所有可能子集的集合具有基数 2^{α}. (提示: 每个子集合 $T\leqslant S$ 定义一个特征函数 $f_T(x)$: $f_T(x)=1$, 当 $x\in T$; $f_T(x)=0$, 当 $x\notin T$.)

4. 证明: 正方形的所有子集的个数等于所有实变量实函数的个数.

*5. (a) 有多少个实数的有限集合?

   (b) 有多少个实数的可数集合?

*6. 有多少个实数的集合, 它们的基数是 $\mathbf{c}$?

7. 约定对一切 $\alpha>0$, 有 $0^0=1$ 和 $0^{\alpha}=0$. 证明: 这个约定与定理 11 的定律(i)～(iv)是相容的.

第十三章 环与理想

§13.1 环

在这一章中，我们将着手研究一般的环以及它们的同态，还要指出后者是如何同理想有关．然后，我们把理想的概念应用到与代数曲线、代数曲面有关的几何上去，并且应用到代数数的分解理论中去(在第十四章中)．我们的基本公设如下所述．

定义 环 A 是这样的元素系统：它在加法运算之下是一个阿贝耳群，在乘法运算之下是封闭的，这个乘法满足结合律，并且对于加法满足分配律．于是，对于环 A 中所有的 a, b, c，有

$$a(bc)=(ab)c,$$
$$a(b+c)=ab+ac, \tag{1}$$
$$(a+b)c=ac+bc.$$

我们还要假定每个环 A 有一个单位元素 $1\neq 0$，满足 $1a=a1=a$，对一切 $a\in A$．

环包括在第一、二、三章中所研究的所有整环和其他交换环，例如 $\mathbf{Z}_m$(模 m 整数)和 $A[x]$, $A[x, y]$(系数在任意给定的交换环 A 中的多项式环)．它还包括非交换环，例如§8.11的四元数环．任意给定域 F 上的所有 $n\times n$ 矩阵组成的集合 $M_n(F)$ 在 $A+B$ 和 AB 运算之下是一个环，当 $n>1$ 时，它也是非交换环．

如果 A 和 B 是任意两个环，那么所有数对 (a, b) (其中 $a\in A$, $b\in B$)组成的集合，在由

$$(a_1, b_1)+(a_2, b_2)=(a_1+a_2, b_1+b_2),$$
$$(a_1, b_1)(a_2, b_2)=(a_1a_2, b_1b_2) \tag{2}$$

定义的两种运算之下是一个环．这样得到的环称为 A 与 B 的*直和*，

记作 $A \oplus B$. 例如，设$\mathbf{Q}$是有理数域，$\mathbf{Z}$是整数环，Q是四元数环，则$\mathbf{Q} \oplus \mathbf{Z} \oplus Q$是一个环. 这个奇妙的例子对很多类型的环给出了某种表示！

交换环理论的大部分可以推广到非交换环上去. 例如，§1.12中给出的环同构的定义，不管是否满足条件$ab=ba$，都适用；§3.3中给出的子环定义也是如此. 此外，交换环中的很多讨论都可以应用到任意环上. 例如我们可以证明，一个环A的子集合S是子环当且仅当$1 \in S$，当b和c在S中时可推出$b-c$和bc都在S中. 另外一些讨论见习题1.

线性代数[①] 矩阵和四元数是具有可加矢量空间结构的一类环的重要例子. 原先这些环是作为比$\mathbf{C}$更广泛的"结合代数系统"来构造的，今天通常称它们为线性结合代数.

定义 *域F上的一个线性代数是一个集合$\mathfrak{A}$，它是F上的一个有限维矢量空间，并且乘法满足结合律和双线性*

$$\alpha(\beta\gamma)=(\alpha\beta)\gamma, \qquad (\text{结合律}) \qquad (3)$$

$$\left.\begin{aligned}\alpha(c\beta+d\gamma)&=c(\alpha\beta)+d(\alpha\gamma),\\(c\alpha+d\beta)\gamma&=c(\alpha\gamma)+d(\beta\gamma),\end{aligned}\right\} \qquad (\text{双线性}) \qquad (4)$$

对F中的所有标量c和d，以及$\mathfrak{A}$中所有元素α, β, γ，上述这些定律都成立. $\mathfrak{A}$的阶是把$\mathfrak{A}$看作矢量空间时它的维数. 当对$\mathfrak{A}$中所有α有$1\alpha=\alpha=\alpha 1$时，$\mathfrak{A}$有一个单位元素1. 如果再增加一个条件：对$\mathfrak{A}$中每个$\alpha \neq 0$，则$\mathfrak{A}$中有元素α^{-1}，满足$\alpha^{-1}\alpha=1$，那么称这个代数是可除代数.

特别是，每个线性代数是一个环.

著名的弗罗比尼乌斯(Frobenius)定理(1878)指出，全体四元数构成实数域上唯一的非交换可除代数.

---

① 关于线性代数这部分材料主要是作为例子的来源，并由于我们对它本身感兴趣，在这里叙述. 它和§13.6可以省略，这并不影响连贯性.

例 1 在实数域上构造一个由"对偶数"组成的代数，它包含两个基元素 δ 和ε，它们是按照法则 $\delta\varepsilon=\varepsilon\delta=\delta$，$\delta^2=0$，$\varepsilon^2=\varepsilon$ 来相乘. 根据这些法则，可以求出 A 的任意两个元素的乘积，因为

$$(a\delta+b\varepsilon)(c\delta+d\varepsilon)=ac\delta^2+ad\delta\varepsilon+bc\varepsilon\delta+bd\varepsilon^2$$
$$=(ad+bc)\delta+bd\varepsilon.$$

可以验证，它满足象乘法结合律那样的必要的公设. 这个例子很象四元数集合，它表明如何通过给出基元素的适当的乘法表可以定义一个代数.

例 2 域 F 上所有 $n\times n$矩阵的全阵代数 $M_n(F)$，以矩阵 E_{ij} 作为它的基底，E_{ij} 的(i,j)位置的元素是 1，其余的元素都是零. 基元素的乘法表是 $E_{ij}E_{jk}=E_{ik}$，$E_{ij}E_{kl}=O(j\neq k)$.

例 3 设 G 是含有元素 $\alpha_1,\cdots,\alpha_n$，并有乘法运算 $\alpha_i\alpha_j=\alpha_k$ 的有限群. 如果 F 是任意域，那么存在 F 上的线性代数$\mathfrak{A}$，它以 G 的元素作为基元素，在 $\mathfrak{A}$ 中，根据 G 的群表，并由双线性来定义乘法，

$$(x_1\alpha_1+\cdots+x_n\alpha_n)(y_1\alpha_1+\cdots+y_n\alpha_n)=\sum_{i,j}(x_iy_j)(\alpha_i\alpha_j).$$

这个代数称为 G 在 F 上的群代数.

特别是，以 α 为生成元的二阶循环群的群代数具有基元素 $1=\alpha^2$ 和 α，并有乘法

$$(x\cdot1+y\alpha)(u\cdot1+v\alpha)=(xu+yv)1+(xv+yu)\alpha.$$

对于基底 $\beta=\frac{1+\alpha}{2}$，$\gamma=\frac{1-\alpha}{2}$，这个群代数有乘法表$\beta^2=\beta$，$\gamma^2=\gamma$，$\beta\gamma=\gamma\beta=0$.

例 4 $2n\times2n$ 矩阵中，在右上方和左下方的 $n\times n$ 矩阵块都是零矩阵的所有矩阵组成的集合构成一个代数，这个代数是 $M_{2n}(F)$ 的一个子环. 它是两个 $M_n(F)$ 的直和.

我们现在证明关于矩阵的类似于凯莱定理（§6.5 定理 8）的

一个定理．同一个域F上的两个代数$\mathfrak{A}$和$\mathfrak{A}'$，当它们的元素之间存在双射$\alpha\longleftrightarrow\alpha'$，并保持三种运算：对一切$\alpha,\beta\in\mathfrak{A}$和一切$c\in F$，有

$$(\alpha+\beta)'=\alpha'+\beta',\ (c\alpha)'=c\alpha',\ (\alpha\beta)'=\alpha'\beta' \tag{5}$$

时，我们称$\mathfrak{A}$与$\mathfrak{A}'$是同构的．

定理 1 具有单位元素的每个n阶线性结合代数与一个$n\times n$矩阵代数同构．

证明 代数$\mathfrak{A}$是元素ξ的矢量空间．同$\mathfrak{A}$中每个元素α相联系的变换T，是通过右乘$\mathfrak{A}$中任意元素ξ（即$\xi T=\xi\alpha$）而得到．因为如(4)中所示的乘法是双线性的，所以T是线性变换．因为现在单位元素是1，由$1\alpha=1\beta$推出$\alpha=\beta$，所以不同的元素α和β导出不同的变换T和U．此外，由代数公设得到

$$\xi(\alpha+\beta)=\xi\alpha+\xi\beta,\qquad \xi(c\alpha)=c(\xi\alpha),\qquad \xi(\alpha\beta)=(\xi\alpha)\beta,$$

所以相应的变换是$\alpha+\beta\longmapsto T+U$，$c\alpha\longmapsto cT$，$\alpha\beta\longmapsto TU$．这意味着对应$\alpha\longmapsto T$是给定的代数与$\mathfrak{A}$上线性变换的代数之间的同构．这些变换又可用矩阵同构地表示出来，因此定理成立．

习　题

1．(a) 证明：在任意环中有$(-a)(-b)=ab$，$-(-a)=a$．

(b) 证明：对所有a，$a0=0a=0$，并且单位元素1是唯一的．

2．证明：由(2)定义的直和实际上是一个环．

3．证明：两个整环的直和不是整环．

4．写出n个已知环的直和的定义，并证明它是一个环．

5．证明：域F上两个线性代数的直和，在适当定义"数乘"运算之后，就可成为F上的线性代数．

6．证明正文中关于描述子环S的命题．

7．证明：线性代数的零元素0，满足$\xi\cdot 0=0=0\cdot\xi$，对一切ξ．

8．"对偶数"的代数是可除代数吗？给出证明．

9. 证明下列各系统是线性代数：

(a) 对所有 α 和 β 满足条件 $\alpha\cdot\beta=0$ 的矢量空间 V_n.

(b) 所有 $n\times n$ 三角形矩阵(对角线下面的元素都为零).

10. 证明：如果 P 是 F 上任意 $n\times n$ 可逆矩阵，那么 $A\mapsto P^{-1}AP$ 是 $M_n(F)$ 的一个自同构. 推广这个结果.

11. 证明：同每个 $n\times n$ 矩阵可交换的 $n\times n$ 矩阵 A 一定是标量矩阵。(提示：A 同每个 E_{ij} 可交换.)

12. 设 $\mathscr{A}$ 是一个代数，证明：$\mathscr{A}$ 中所有那些与 $\mathscr{A}$ 的每个元素可交换的元素 z 组成的集合 $\mathscr{S}$ 是 $\mathscr{A}$ 的子代数.（它称为 $\mathscr{A}$ 的中心.）

§13.2 同　态

给定两个环 A 和 A' 以及对应 $a\longmapsto aH$，如果对 A 中每个元素 a, aH 是 A' 中唯一确定的元素，并且对 A 中一切元素 a, b, 有

$$(a+b)H=aH+bH,\ (ab)H=(aH)(bH),\ 1H=1',\qquad (6)$$

则我们称对应 $a\longmapsto aH$ 是 A 到 A' 的同态. 简单地说，正如§3.3中交换环的情形一样，同态是一个保持单位元素，保持和与积的映射. 同群的情形一样，映上的同态也称为满同态.

从环 A 到 A' 的同态 H 一定是由 A 的加法群到 A' 的加法群的同态，所以 H 具有§6.11中对于群已经证明了的性质

$$0H=0',\ (-a)H=-(aH),\ (a-b)H=aH-bH,\qquad (7)$$

这里 $0'$ 是环 A' 的零元素，即 A' 的加法群的单位元素.

我们所熟悉的把每个整数 a 映射到模 m 剩余类的对应 $a\longmapsto a_m$ 是整数环 $\mathbf{Z}$ 到 $\mathbf{Z}_m$ 的同态. 如果 $f(x)$ 是系数在整环 D 中的任意多项式，那么用 D 中固定元素 b "替换" $f(x)$ 中的 x 而得到的对应 $f(x)\longmapsto f(b)$ 是多项式整环 $D[x]$ 到 D 的一个同态，因为未定元 x 的多项式形式的加法法则和乘法法则，对于相应的 b 的多项式表达式，当然适用. 如果 $\mathbf{Q}[x]$ 是有理系数的多项式环，那么对应 $f(x)\longmapsto f(\sqrt{2})$ 是多项式环 $\mathbf{Q}[x]$ 到由所有数 $a+b\sqrt{2}$ 构成的域上

的满同态(见§2.1的讨论). 两个环A与B的直和$A\oplus B$, 通过对应 $(a, b)\longmapsto b$ 满同态地映到被加项B; 正好根据直和运算的定义(2), 这个对应保持和与积.

为了明确地描述一个具体的同态, 我们自然要问, 什么时候第一个环中两个元素a和b在第二个环中有相同的象. 根据法则(7), 只有当它们的差的象$(a-b)H=0'$时, 才可能发生这种情况. 因此我们寻求这样元素的集合, 这些元素通过H映射到A'的零元素$0'$上. 例如, 同态$\mathbf{Z}\longrightarrow\mathbf{Z}_m$把模$m$的所有倍数$km$映上到零. 所有这些倍数的集合在减法之下是封闭的, 还有, 这个集合的元素同$\mathbf{Z}$的任何元素相乘后仍在这个集合中. 类似地, 同态$f(x)\longmapsto f(b)$把所有可被$x-b$整除的多项式映射到零, 再没有其他多项式可以映射到零. 所有这些多项式组成的集合S在减法和用$D[x]$的所有元素(不管这些元素是否在S中)与它相乘的运算之下是封闭的. 这两个例子蕴含着下面的定义和定理(参看§3.8).

定义 环A中的理想C是具有下述性质的A中非空集合:

(i) c_1和c_2在C中, 可推出c_1-c_2在C中;

(ii) c在C中, a在A中, 可推出ac和ca都在C中.

定理2 在环A的任意同态H中, 由所有映射到零的元素组成的集合是A中的一个理想.

为了在一般情形下证明定理2, 设C是A中所有满足$cH=0'$的元素c的集合, 其中$0'$是象A'的零元素. 那么, 对A中的任意元素a, 有$(ac)H=(aH)(cH)=(aH)0'=0'$和$(ca)H=(cH)(aH)=0'$, 这就证明了性质(ii). 此外, 根据(7), 由$c_1H=c_2H=0'$, 可得到

$$(c_1-c_2)H=c_1H-c_2H=0'-0'=0',$$

因此证明了性质(i).

这个结果表明, 环中的理想类似于群中的正规子群. 为表示

这种类似，我们称通过同态H映射到零的所有元素组成的集合为H的核。当H是满射(即满同态)时，我们称环B是环A在同态H之下的满同态象，所以每个元素$b\in B$是某个$a\in A$在H之下的象aH.

定理3 环A的满同态象由它的核确定(除同构外).

证明 我们必须证明，如果H和K分别是A到A'和A''上的满同态，并且$aH=0'$当且仅当$aK=0''$，那么A'和A''是同构的．很自然地，设元素$a'\in A'$对应于$a''\in A''$当且仅当这两个元素在A中有公共的象源a，所以对某个a，当$aH=a'$, $aK=a''$时，有$a'\longleftrightarrow a''$. 这个对应是一对一的：在这个对应下，对$A'$中每个$a'$，在$A''$中有一个且只有一个$a''$与$a'$对应．为证明这一点，首先注意，对$A'$中每个$a'$，在$A$中至少有一个象源$a$，因此在$A''$中至少有一个$a''=aK$与$a'$对应．其次，如果$a'\longleftrightarrow a''$, $a'\longleftrightarrow b''$，那么对$A$中某两个$a, b$，有

$$aH=a',\quad aK=a'',\quad bH=a',\quad bK=b'',$$

因此$(a-b)H=a'-a'=0'$，根据假设可推出$0''=(a-b)K=a''-b''$.这个对应也保持和与积，这是因为如果$a'\longleftrightarrow a''$, $b'\longleftrightarrow b''$，那么

$$a'+b'=(a+b)H\longleftrightarrow(a+b)K=a''+b'',$$

$$a'b'=(ab)H\longleftrightarrow(ab)K=a''b'',$$

这里a是a'和a''的公共象源，b是b'和b''的公共象源．

理想的两个性质(i)和(ii)有几个直接推论．任意理想C包含某元素c，因此(i)表明$c-c=0$在C中，所以对C中任意c, $0-c=-c$也在C中．根据性质(i)，我们得到，C中任意两个元素之和$c_1+c_2=c_1-(-c_2)$也在C中。于是，因为$1\in A$，所以A的非空子集C是A的理想当且仅当任意线性组合$a_1c_1\pm a_2c_2$和$c_1a_1\pm c_2a_2$在C中，其中c_1和c_2在C中，a_1和a_2在A中．特别是，A的理想不一定是A的子环，因为它可以不包含A的单位元素．整个环A

和仅由一个元素 0 组成的 $\{0\}$ 总是任意环 A 的理想．它们称为环 A 的假理想．任意其他理想称为真理想．相应地，环 A 的真满同态是这样一个同态，它的核是 A 的真理想，所以真满同态不是一个同构（同构只把 $\{0\}$ 映到 $0'$）．

定理 4 可除环没有真满同态象．

证明 只须证明可除环 D 没有真理想．设 C 是 D 中任意理想，它不是理想 $\{0\}$，于是它包含一个元素 $c\neq 0$．由性质(ii)，C 包含 $1=c^{-1}c$，再由性质(ii)，C 包含整个可除环的任意元素 $a=a\cdot 1$．因此 C 不是真理想，如断言所述．

证毕

如果 b 是交换环 A 的元素，那么 b 的所有倍数 xb（其中变数 $x\in A$）的集合 (b) 是一个理想，因为可以验证性质(i)和(ii)．这个理想 (b) 称为主理想，它是 A 中包含 b 的最小理想．我们回忆一下，由 § 1.7 定理 6 可知，整数环 $\mathbf{Z}$ 中的每个理想都是主理想．根据 § 3.8 定理 11，任意域 F 上的一个未定元的多项式整环 $F[x]$ 中，上述结论同样成立．

在两个变量的有理系数多项式环 $\mathrm{Q}[x,y]$ 中，常数项为零的所有多项式的集合 C 是一个理想．它并不是主理想，因为两个多项式 x 和 y 虽然都在 C 中，但它们不能都是其中一个多项式的倍式，也不能同时是同一个多项式 $f(x,y)$ 的倍式．虽然这个理想 C 不能由任意单个多项式 $f(x,y)$ 生成，但是它的所有元素都可用含有多项式系数的线性组合 $xg(x,y)+yh(x,y)$ 来表示，所以整个理想是通过两个生成元 x 和 y 的线性组合给出．

现在考虑由交换环 A 中任意给定的有限个元素的集合生成的理想．如果一个理想 C 包含元素 $c_1, c_2, \cdots, c_m$，那么它必包含所有线性组合 $\sum x_i c_i$，其中系数 x_i 在 A 中．而集合

$$(c_1, c_2, \cdots, c_m) = [\text{所有元素} \sum_i x_i c_i, \quad x_i \in A] \tag{8}$$

本身是一个理想，这是因为

$$\sum_i x_i c_i - \sum_i y_i c_i = \sum_i (x_i - y_i) c_i,$$

$$a\left(\sum_i x_i c_i\right) = \sum_i (a x_i) c_i.$$

也就是说，这个集合具有关于理想所要求的性质(i)和(ii)．因为A具有单位元素1，所以每个元素c_i必然是集合(8)中的一个元素

$$c_i = 0 \cdot c_1 + \cdots + 0 \cdot c_{i-1} + 1 \cdot c_i + 0 \cdot c_{i+1} + \cdots + 0 \cdot c_m.$$

因此，由(8)式定义的集合$(c_1, \cdots, c_m)$是包含$c_1, \cdots, c_m$的A的一个理想，并且它还被包含在包含所有c_i的任何理想中，它称为以$c_1, \cdots, c_m$为基底的理想．(这种基元素不能同矢量空间的基底相类比，因为$x_1 c_1 + \cdots + x_m c_m = 0$不一定能推出$c_1 = \cdots = c_m = 0$.)

在大多数熟悉的整环中，每个理想都有一组有限基底，但是也存在一些整环情况并非如此．

习　题

1．在下列映射中，哪些是同态，为什么？对于是同态的映射，描述映射到零的理想．

(a) $a \mapsto 5a$，a是$\mathbf{Z}$中整数．

(b) $f(x) \mapsto f(\omega)$，$f(x)$是$\mathbf{Q}[x]$中多项式，ω是三次单位根．

(c) $f(x, y) \mapsto f(t, t)$，是$F[x, y]$到$F[t]$的映射(x, y, t是未定元)．

2．证明：交换环的每个同态象是可交换的．

3．在多项式$f(x, y) = a + b_1 x + b_2 y + c_1 x^2 + c_2 xy + c_3 y^2 + \cdots$组成的环$\mathbf{Q}[x, y]$中，下列多项式集合中哪些是理想？如果那个集合是理想，就求出它的基底．

(a) 常数项为零$(a=0)$的所有多项式$f(x, y)$．

(b) 不含x的所有多项式$f(x, y)$ $(b_1 = c_1 = c_2 = \cdots = 0)$．

(c) 不含二次项的所有多项式($c_1=c_2=c_3=0$).

4. (a) 求出 $\mathbf{Z}_6$ 中的所有理想.

(b) 求出 $\mathbf{Z}_6$ 的所有同态象.

5. 详细证明: 第一章中定义的环 $\mathbf{Z}/(m)=\mathbf{Z}_m$ 是 $\mathbf{Z}$ 的唯一的真满同态象.

6. (a) 对每个 m 求出 $\mathbf{Z}_m$ 中的所有理想.

(b) 求出 $\mathbf{Z}_m$ 的所有满同态象.

7. 求出两个域的直和中的所有理想. 并推广这个结果.

8. 求出直和 $\mathbf{Z}\oplus\mathbf{Z}$ 中的所有理想, 这里 $\mathbf{Z}$ 是整数环.

*9. 证明: 如果 C_1 和 C_2 分别是环 A_1 和 A_2 中的理想, 那么直和 $C_1\oplus C_2$ 是直和 $A_1\oplus A_2$ 中的理想, 并且这个直和中的每个理想都具有这种形式.

10. 证明: 在整环中, $(a)=(b)$ 当且仅当 a 和 b 是相伴(§ 3.6).

11. 证明: 如果 A 是一个交换环, 其中每个理想都是主理想, 那么 A 中任意两个元素 a 和 b 都有最大公因子 d, 并可表成 $d=ra+sb$.

*12. 设 A 是包含域 F 的环, A 和 F 具有同一个单位元素(例如 A 可以是域 F 上的多项式环). 证明: A 的每个真同态象包含一个与 F 同构的子域.

*13. 设 $\mathbf{Z}_p^*$ 是由所有有理数 $\frac{m}{n}$(其中分母与给定的素数 p 互素)组成的环. 证明: $\mathbf{Z}_p^*$ 中每个真理想具有形式 (p^k), 其中 k 是某正整数.

§ 13.3 商　　环

对环的每个同态, 都存在对应的理想, 该理想的元素在同态下映射到零. 反过来, 给定一个理想, 我们现在将构造一个相应的同态象. 环 A 中的理想 C 是 A 的加法群的子群. A 中每个元素 a 属于一个陪集, 这个陪集常常称为剩余类 $a'=a+C$, 它是由所有和 $a+c$(变量 $c\in C$)组成. 两个元素 a_1 和 a_2 属于同一个陪集当且仅当它们的差在这个理想 C 中. 因为加法是可交换的, 所以 C 是加法群 A 的正规子群, 因此 C 的所有陪集构成阿贝耳商群. 在这个商群中, 两个陪集的和是另一个陪集, 它是通过把两个代表元相加而得到, 即

$$(a_1+C)+(a_2+C)=(a_1+a_2)+C. \tag{9}$$

§ 6.13 中已经证明，这个和不依赖于在给定的陪集中元素 a_1 和 a_2 的选择.

为了构造两个陪集的乘积，在第一个陪集中选取任意元素 a_1+c_1，在第二个陪集中选取任意元素 a_2+c_2. 乘积

$$(a_1+c_1)(a_2+c_2)=a_1a_2-(a_1c_2+c_1a_2+c_1c_2)=a_1a_2+c'$$

总是陪集 a_1a_2+C 中的一个元素，因为根据理想的性质(ii)，a_1c_2，c_1a_2，c_1c_2 这些项都在理想 C 中. 因此，第一个陪集中的元素与第二个陪集中的元素的所有乘积都在同一个陪集中，这个乘积陪集是

$$(a_1+C)(a_2+C)=a_1a_2+C. \tag{10}$$

陪集乘法的结合律和分配律立即从 A 中相应的定律可以得到. 包含 1 的陪集起单位元素的作用，所以 A 中 C 的全体陪集构成一个环.

正是根据陪集运算的定义(9)和(10)，把 A 的每个元素映射到它的陪集的对应 $a\longmapsto a'=a+C$ 是一个满同态. 在这个满同态象中，零元素是陪集 $0+C$，所以 C 的元素都被映射到零. 这些结果可以总结如下：

定理 5 在定义(9)和(10)之下，环 A 中任意理想 C 的全体陪集构成一个环，称为商环① A/C. 函数 $a\longmapsto a+C$ 把 A 的每个元素映射到包含它的陪集，它是 A 到商环 A/C 的一个满同态，而且这个满同态核是给定的理想 C.

推论 1 如果 A 是可交换的，那么 A/C 也是可交换的.

理想与同态的关系现在已经齐全了. 特别是，定理 3 关于唯一性的断言可以改述如下：

---

① 环 A/C 也常被称为剩余类环，因为它的元素是 C 在 A 中的剩余类(陪集).

推论 2 如果满同态 H 把 A 映射到 A'，并有同态核 C，那么 A' 与商环 A/C 同构.

模 m 整数的环 $\mathbf{Z}_m$ 现在可以描述成商环 $\mathbf{Z}/(m)$. 反过来，由这个例子所启发，当 $(a-b)\in C$ 时，我们常常写成 $a\equiv b(C)$，并且说 a 和 b 是同余的，模环 R 的理想 C.

商环的每个性质都反映在它的生成理想 C 的相应的性质中. 为了解释这个原理，我们定义极大理想和素理想. 如果 A 中包含理想 C 的理想只能是 C 和 A 本身，则我们称 $C<A$ 是极大理想①. 如果 A 中理想 P 包含乘积 ab 时，至少包含其中一个因子 a 或 b，则我们称 P 为素理想.

在交换环中，素理想起着特殊的作用. 例如，在整数环 $\mathbf{Z}$ 中，主理想 (p) 是素理想当且仅当 p 是素数，因为当 p 是素数，而不是其他情形时，两个整数的乘积 ab 是 p 的倍数当且仅当有一个因子是 p 的倍数.

定理 6 如果 A 是交换环，那么商环 A/C 是整环当且仅当 C 是素理想. 商环 A/C 是域当且仅当 C 是 A 的极大理想.

证明 交换环 A/C 是整环当且仅当它没有零因子（§1.2定理 1）. 这个条件用公式写成

$$a'b'=0, \text{仅当 } a'=0 \text{ 或 } b'=0, \tag{11}$$

这里 a' 和 b' 分别是元素 a 和 b 在 A 中的陪集. 现在 C 的陪集 a' 是零当且仅当 a 在理想 C 中，则上述条件可以改写成

$$ab \text{ 在 } C \text{ 中，仅当 } a \text{ 在 } C \text{ 中或 } b \text{ 在 } C \text{ 中}. \tag{12}$$

这恰好就是素理想 C 的定义.

其次，假定 C 是极大理想，并设 b 是 A 中不属于 C 的任意元素. 那么可以证明，所有元素 $c+bx$（其中任意 $c\in C$，任意 $x\in A$）组

① "极大理想"有时用"无因子理想"代替.

成的集合是一个理想．这个理想包含C，并包含不在C中的元素b．因为C是极大理想，所以这个理想一定是整个环A．特别是，单位元素1在这个理想中，所以对某个a，有$1=c+ba$．用陪集来表示，这个方程写成$1'=b'a'$．于是，对任意陪集$b'=b+C\neq C$，我们已求出互逆陪集$a'=a+C$，这就是说，陪集组成的交换环是一个域．反过来，如果A/C是一个域，我们可以证明C是极大理想（习题10）． 证毕

因为每个域是一个整环，所以定理6意味着每个极大理想是素理想．然而反过来，素理想不一定是极大理想．例如，考虑同态$f(x,y)\longmapsto f(0,y)$，它把系数在域F上的未定元x和y的多项式整环$F[x,y]$映射到较小的整环$F[y]$上．因此映射到零的理想是主理想(x)，它是由x的所有倍式（多项式）组成．因为象环$F[y]$实际上是一个整环，所以这个理想(x)是一个素理想，这也可以直接验证．但是$F[y]$不是域，所以(x)不可能是极大理想．实际上，它包含在较大的理想(x,y)之中，(x,y)是由常数项为零的所有多项式组成．

习　题

1. 证明：陪集的乘法满足结合律和分配律．

2. 设模一个理想$C\leqslant A$的同余定义为：$a\equiv b(\mathrm{mod}C)$当且仅当$a-b$在C中．证明：同余可以相加和相乘．C的陪集是由相互同余的元素组成．

3. 详细证明定理5的推论1．

4. 求出整数环$\mathbf{Z}$中所有素理想．

5. 求出域F上多项式环$F[x]$中的所有素理想和所有极大理想．

*6. 不用定理6证明：整环中每个极大理想都是素理想．

*7. 在整系数多项式整环$\mathbf{Z}[x]$中，求出不是极大理想的素理想．

8. 证明：在所有数$a+b\omega$（a,b为整数，ω是虚三次单位根）组成的整环$\mathbf{Z}[\omega]$中，(2)是素理想．并描述商环$\mathbf{Z}[\omega]/(2)$．

9. 在多项式环 $\mathbf{Q}[x, y]$ 中，下列理想中哪些是素理想？哪些是极大理想？

(a) (x^2),　　(b) $(x-2, y-3)$,

(c) $(y-3)$,　　(d) (x^2+1),

(e) (x^2-1),　　(f) $(x^2+1, y-3)$.

10. 证明：如果商环 A/C 是一个域，那么 C 是极大理想.

11. 求出一个熟悉的环与下面每个商环 A/C 同构：

(a) $A=\mathbf{Q}[x]$, $C=(x-2)$;

(b) $A=\mathbf{Q}[x]$, $C=(x^2+1)$;

(c) $A=\mathbf{Q}[x,y]$, $C=(x,y-1)$;

(d) $A=\mathbf{Z}[x]$, $C=(3,x)$;

(e) $A=\mathbf{Z}_p^*$, $C=(p)$, 这里 $\mathbf{Z}_p^*$ 的定义见 § 13.2 的习题 13.

12. (第二同构定理)设 $C>D$ 是环 A 中的两个理想.

(a) 证明：商 C/D 是 A/D 中的理想.

(b) 证明：A/C 与 $(A/D)/(C/D)$ 同构. (提示：两个同态的乘积是一个同态.)

*§ 13.4 理想的代数

理想之间的包含同数之间的整除性有着密切的关系. 在整数环 $\mathbf{Z}$ 中, $n|m$ 意味着 $m=an$, 因此 m 的每个倍数是 n 的倍数. n 的倍数组成主理想 (n), 所以条件 $n|m$ 意味着 (m) 包含在 (n) 中. 反过来, $(m)\subset(n)$ 特别意味着 m 在 (n) 中, 因此 $m=an$. 所以

$$(m)\subset(n) \quad \text{当且仅当} \quad n|m.$$

更一般地, 在任意交换环 R 中, $(b)\subset(a)$ 可推出, 对某个 $x\in R$ 有 $b=ax$, 即 $a|b$. 反过来, 如果 $a|b$, 那么对某个 $x\in R$ 有 $b=ax$, 于是对所有的 $by\in(b)$, 有 $by=axy\in(a)$, 因此 $(b)\subset(a)$. 这就证明了

定理 7 在任意交换环 R 中,

$$(b)\subset(a) \quad \text{当且仅当} \quad a|b. \tag{13}$$

但要注意: "较大"的数对应着"较小"的理想; 例如, 6 的所有倍数

组成的理想(6)真正地包含在所有偶数组成的理想(2)中.

最大公因子和最小公倍数在理想理论中也有相应的解释. 整数 n 和 k 的最小公倍数 m 是 n 和 k 的倍数, 并且是 n 和 k 的其他每个公倍数的因子. 于是, m 的所有倍数的集合 (m) 是 n 和 k 的所有公倍数的集合, 刚好也是主理想 (n) 和 (k) 的公共元素组成的集合. 这个情况可以推广到任意环(不一定是交换环)的任意理想上, 如下所述.

可以证明, 环 A 的任意两个理想 B 和 C 的交 $B\cap C$ 是一个理想. 设 D 是 A 的任意其他理想, 则理想 $B\cap C$ 具有三个性质:

(i) $B\cap C\subset B$,

(ii) $B\cap C\subset C$,

(iii) 由 $D\subset B$ 和 $D\subset C$ 可推出 $D\subset B\cap C$. 于是在格论意义之下, 这个交是 B 和 C 的最大下界.

对偶于交的是两个理想的和. 如果 B 与 C 是 A 中两个理想, 我们则可验证集合

$$B+C=[\text{所有的和 } b+c, b\in B, c\in C] \tag{14}$$

是 A 中一个理想. 因为包含 B 和 C 的任意理想一定包含所有的和 $b+c$, 所以这个理想 $B+C$ 包含 B 和 C, 并且包含在每个包含 B 和 C 的理想中. 于是在格论意义下 $B+C$ 是 B 与 C 的最小上界也是 B 与 C 的并。

定理 8 由(14)式的和 $B+C$ 给出的并与由 $B\cap C$ 给出的交, 在通常包含关系之下, 环 A 中的全体理想构成一个格.

如果整数 m 和 n 有最大公因子 d, 那么理想之和 $(m)+(n)$ 恰好是主理想 (d). 这是因为根据(13), 有 $(d)\supset(m)$ 和 $(d)\supset(n)$; 由于 d 有表达式 $d=rm+sn$, 所以包含 m 和 n 的任意理想一定包含 d, 因而也包含 (d) 的所有元素. 因此, (d) 是 (m) 和 (n) 的并, 即 $(d)=(m)+(n)$.

前面的研究可以推广如下：

引理 *在交换环 R 中，两个主理想的和 $(b)+(c)$ 本身是主理想 (d) 当且仅当 d 是 b 和 c 的最大公因子.*

我们把引理的证明留给读者.

一般地，在交换环中，如果理想 B 和 C 是由如下的基底生成的：

$$B=(b_1,\cdots,b_m),\quad C=(c_1,\cdots,c_n) \tag{15}$$

那么我们有，对任意 $b+c\in B+C$，按(8)式得

$$b+c=\sum_i x_i b_i+\sum_j y_j c_j$$

也就是说，$B+C$ 是由 $b_1,\cdots,b_m$ 和 $c_1,\cdots,c_n$ 生成，所以

$$(b_1,\cdots,b_m)+(c_1,\cdots,c_n)=(b_1,\cdots,b_m,c_1,\cdots,c_n). \tag{16}$$

这个法则同基底的自然变换结合起来，可以用来明显地计算出整数的最大公因子. 例如，

$$\begin{aligned}(336)+(270)&=(336,270)=(336-270,270)=(66,270)\\&=(66,270-4\times 66)=(66,6)=(6),\end{aligned}$$

所以 336 和 270 的最大公因子是 6.

在任意交换环中，我们还可以定义任意两个理想 B 与 C 的乘积 BC，

$$BC=[\text{所有和 } b_1c_1+\cdots+b_mc_n,\ b_i\in B,\ c_j\in C]. \tag{17}$$

实际上，这个集合是一个理想，它是由所有的乘积 bc 生成的，其中因子 b 在 B 中，因子 c 在 C 中，所以它也是包含所有这些乘积的最小理想. 特别是，两个主理想 (b) 和 (c) 的乘积就是由已知元素 b 和 c 的乘积生成的主理想 (bc). 更一般地，如果理想 B 和 C 是由(15)式的基底确定的，任意乘积 bc 具有形式

$$bc=\left(\sum_i x_i b_i\right)\left(\sum_j y_j c_j\right)=\sum_{i,j}(x_i y_j)(b_i c_j),$$

因此乘积理想 BC 有基底

$$BC=(b_1c_1, b_1c_2, \cdots, b_mc_{n-1}, b_mc_n). \tag{18}$$

这种乘积对于代数数论(§14.10)是有用的.

习　　题

1. 详细证明: $B\cap C$ 和 $B+C$ 总是理想.

2. 证明: (17)式的乘积 BC 是一个理想.

3. 画出 $\mathbf{Z}_{24}$ 中所有理想的格图.

4. 设 $f(x)$和 $g(x)$是域上的多项式, $d(x)$是它们的最大公因式. 证明: $(f(x)+g(x))=(d(x))$.

5. 用理想基底的方法计算最大公因子(280, 396)和(8624, 12825).

6. 证明: 整数环 $\mathbf{Z}$ 中每个理想可唯一地表示成素理想的乘积.

7. 证明下列变换理想的基底的法则:

$$(c_1, c_2, \cdots, c_m)=(c_1+xc_2, c_2, \cdots, c_m),$$
$$(xc_1, c_2, \cdots, c_m)=(c_1, c_2, \cdots, c_m).$$

8. 在 $R[x, y]$中, 化简下面理想的基底:

$(x^3+y,\ 3y,\ 4x^4+x^2)$,

$(x^2+3xy+y^2,\ 2x^2-y^2,\ x^2+6xy,\ x^3+y^2)$.

9. (a) 证明: 在任意交换环中, $BC\subset B\cap C$.

(b) 给出一个例子说明 $BC<B\cap C$ 是可能的.

(c) 证明: $B(C+D)=BC+BD$.

*10. 证明: 任意环中所有理想构成的格是一个模 (按§11.7 习题 10 的定义).

11. 在交换环 A 中, 设 $B:C$ 表示满足 $xc\in B(c\in C)$的所有元素 x 组成的集合.

(a) 证明: 如果 B 和 C 是理想, 那么 $B:C$ 也是 A 中的理想(称为"理想商").

(b) 证明: $(B_1\cap B_2):C=(B_1:C)\cap(B_2:C)$.

(c) 证明: $B:C$ 是满足 $CX\subset B$ 的所有理想 X 的最小上界(并).

12. 证明: 如果环 R 包含理想 B 和 C, 并满足 $B\cap C=0$, $B+C=R$, 那么 R 同构于 B 与 C 的直和.

§13.5 多项式理想

理想的概念在现代代数几何中是基本的. 当我们研究三维空间的代数曲线时，由于引入理想的概念，使推理立即变得明显了.

一般在 n 维矢量空间 F^n 中，一个(仿射) 代数簇定义为点集合 V, 这个集合中的所有点$(x_1, \cdots, x_n)$都满足有限个适当的多项式方程的方程组

$$f_1(x_1, \cdots, x_n)=0, \cdots, f_m(x_1, \cdots, x_n)=0. \tag{19}$$

例如, 在 $\mathbf{R}^3$ 中，平行于 xy 平面且在 xy 平面上方两个单位的平面中的圆 C (圆心在 z 轴上，半径为2) 通常可解析地描述为空间中满足联立方程

$$x^2+y^2-4=0, \quad z-2=0 \tag{20}$$

的点 (x, y, z) 的集合. 这些方程把曲线 C 描绘成圆柱面与一平面的交线. 但是也可以用等价的联立方程

$$x^2+y^2+z^2-8=0, z-2=0, \tag{21}$$

以同样的精确度把 C 描绘成一个球面与平面 $z=2$ 的交线. 还可能有其他的描述, 例如用方程组

$$x^2+y^2-4=0, \quad x^2+y^2-2z=0. \tag{22}$$

这些方程把 C 描绘成圆柱面与旋转抛物面 $x^2+y^2=2z$ 的交线.

我们可以通过用全体这样的多项式方程 (即曲线 C 的点所适合的方程)描绘 C 来避免上述的含糊. 如果 $f(x, y, z)$和 $g(x, y, z)$ 是任意两个多项式, 它们的值在 C 上恒等于零, 那么它们的和与差在 C 上也恒等于零. 同样地, 不管什么样的多项式 $a(x, y, z)$ 与 $f(x, y, z)$的乘积 $a(x, y, z)f(x, y, z)$在 C 上也恒等于零. 这就意味着, 在 C 上的值恒等于零的所有多项式组成的集合是一个理想. 那么这个理想(而不是它的任意一对个别的元素)就是 C 的根本的

描述．我们现在将证明所有这样的方程组成的集合是一个理想．

定理9 在 F^n 中，在给定的集合 S 上恒等于零的所有多项式的集合 $J(S)$ 是 $F[x_1, \cdots, x_n]$ 中的一个理想．

因为如果 $p(x_1, \cdots, x_n)$ 在给定的点上等于零，那么 p 的所有倍式在该点上也等于零，而当 p 和 q 在给定的点上等于零，则 $p\pm q$ 在该点上也等于零．对于在给定的集合 S 上恒等于零的多项式，上述结论同样正确．实际上，$J(S)$ 刚好是在不同的点 $\xi\in S$ 上等于零的多项式理想 $J(\xi)$ 的交．

例如，上面所讨论的圆 C 的情形中，$J(C)$ 是由所有线性组合

$$h(x, y, z)=a(x, y, z)(x^2+y^2-4)+b(x, y, z)(z-2) \qquad (23)$$

构成的理想，其中系数为多项式 $a(x, y, z)$ 和 $b(x, y, z)$．这就是说，$J(C)$ 就是以 x^2+y^2-4 和 $z-2$ 为基底的理想 $(x^2+y^2-4, z-2)$．(21)中的多项式生成同一个理想，因为这些多项式是(20)的那两个多项式的线性组合，而反过来，(20)的多项式可以由(21)的两个多项式的线性组合而得到．于是由这条曲线确定的多项式理想有各种各样的基底，

$$\begin{aligned}(x^2+y^2-4, z-2)&=(x^2+y^2+z^2-8, z-2)\\&=(x^2+y^2-2z, z-2).\end{aligned} \qquad (24)$$

商环 $\mathbf{R}[x, y, z]/(x^2+y^2-4, z-2)$ 有一个重要意义．即它与定义在 C 上的所有函数组成的环同构(参看§3.2)，这些函数可定义为变量 x, y, z 的多项式．显然，它与 $\mathbf{R}[x, y]/(x^2+y^2-1)$ 同构，因此用通常的同化规则可知，它与所有三角多项式 $p(\cos\theta, \sin\theta)$ 构成的环同构．这个商环称为 C 上多项式函数环，由它扩张成的域称为 C 上有理函数域．

三次挠线 C_3: $x=t, y=t^2, z=t^3$ 是一条代数曲线，它(不象 C 那样)可以用参数 t 的多项式函数来定义．显然，给定的点 (x, y, z) 在 C_3 上当且仅当 $y=x^2$，$z=x^3$．因此 C_3 是 $\mathbf{R}^3$ 中由理想 $M=$

$(y-x^2, z-x^3)$定义的代数曲线.

根据定义，多项式 $p(x, y, z)$ 在 C_3 上恒等于零当且仅当对所有 $t\in\mathbf{R}, p(t, t^2, t^3)=0$. 现在考虑同态①

$$f(x, y, z)\longmapsto f(t, t^2, t^3) \quad (t \text{ 是未定元}). \tag{25}$$

显然，对 C_3 上的所有点，有 $y=x^2, z=x^3$，这表明 $y-x^2$ 和 $z-x^3$ 将位于我们的理想M之中. 但是反过来，我们注意，变量替换 $y=y'+x^2, z=z'+x^3$ 将把任意多项式 $f(x, y, z)$变为 $f'(x, y', z')$，并注意，按照这种形式，同态(25)是

$$f'(x, y', z')\longmapsto f'(t, 0, 0). \tag{25'}$$

这个对应把 f' 中包含 y' 或 z' 的项映射到零，而不是别的，所以被映射到零的多项式就是线性组合 $g(x, y, z)y'+h(x, y, z)z'$. 因此，我们的理想$M$恰是以 $y'=y-x^2, z'=z-x^3$ 为基底的理想$(y', z')=(y-x^2, z-x^3)$. 这就把曲线 C_3 表示成一个抛物柱面与另一个柱面的交线. 在 C_3 的进一步分析中，商环 $\mathbf{R}[x, y, z]/M$ 起着重要作用. 映射(25)表明这个商环与多项式环 $\mathbf{R}[t]$同构.

两个理想的和有简单的几何解释. 例如，在 $\mathbf{R}[x, y, z]$中，主理想$(z-2)$表示平面 $z=2$，因为这个理想中所有的多项式 $f(x, y, z)(z-2)$，当其中的 x, y, z 用平面 $z=2$ 上点的坐标代替时，都恒等于零. 类似地，主理想(x^2+y^2-4)定义了一个以 z 轴为轴的半径为 2 的圆柱面. 根据法则(16)，这两个理想的和是$(x^2+y^2-4, z-2)$. 我们刚刚看到，这个和(23)表示一个圆，它是平面和圆柱面的交线. 事实上显然有：两个理想的和所对应的轨迹是各理想所确定的轨迹的交.

反过来，多项式环 $\mathbf{R}[x_1, \cdots, x_n]$ 中的任意理想 J 确定一个相

---

① 注意，(25)式定义一个同态这一事实并不显然，证明它需要把 §3.1 的定理 1 推广.

应的轨迹，这个轨迹是由 n 维空间中使得对每个多项式 $f\in J$ 有 $f(a_1, \cdots, a_n)=0$ 的所有的点 $(a_1, \cdots, a_n)$ 组成. 希尔伯特基底定理断言: J 具有有限基底 $f_1, \cdots, f_m$, 所以相应的轨迹 V 的确是一个代数簇. 可是，这个簇的理想 $J(V)$ 可以大于给定的理想 J(参看下面的习题 3).

习　题

1. 求 $\mathbf{R}^3$ 中具有参数方程 $x=t+1$, $y=t^3$, $z=t^4+t^2$ 的曲线所对应的理想.

2. 证明: 由两个线性无关的线性多项式生成的任意理想 $(ax+by+cz,\ a'x+b'y+c'z)$ 确定 $\mathbf{R}^3$ 中的一条直线.

3. (a) 证明: 在 $\mathbf{R}[x, y, z]$ 中的理想 (x, y) 与 $(x^2,\ xy,\ y^2)$ 确定相同的代数簇.

(b) 证明: 任意理想和它的平方确定相同的轨迹.

4. 详细证明: 在任意轨迹 C 上恒等于零的 $\mathbf{R}[x_1, \cdots, x_n]$ 中多项式的集合是一个理想.

5. (a) 在三维空间中, $xy=0$ 所确定的轨迹是什么?

(b) 证明: 由两个主理想的乘积所确定的轨迹是各主理想所确定的轨迹的并.

(c) 把(b) 的结果推广到任意理想上. (提示: 如果在乘积的轨迹中有一点不在第一个因子所确定的轨迹上, 那么在第一个理想中至少有一个多项式在这点不为零.)

(d) 两个理想的交所确定的轨迹是什么?

6. (a) 计算下面"双有理"变换的逆:

$$T:\ x'=x,\ \ y'=y-x^2,\ \ z'=y-x^3$$

(b) 证明: 所有形为 $x'=x$, $y'=y+p(x)$, $z=z'+q(x, y)$ $(p, q$ 是多项式)的变量替换组成的集合是一个群.

(c) 证明: 每个这样的替换诱导出一个环 $\mathbf{R}[x, y, z]$ 上的自同构.

7. (a) 证明: 如果 H 是交换环 A 中一个理想, H 的根式是由 A 中所有使得某个幂 $x^m\in H$ 的 x 所组成的集合 $\sqrt{H}$, 那么 $\sqrt{H}$ 是一个理想.

(b) 证明：如果H是多项式环$\mathbf{C}[x,y,z]$中的一个理想，V是相应的轨迹，那么$J(V)$包含$\sqrt{H}$.（提示：希尔伯特零点定理断言$J(V)=\sqrt{H}$.）

8. 描述下面两种情形下，由$x^2+y^2=0$所确定的轨迹。

(a) 在$\mathbf{R}^3$中.　　　　　　(b) 在$\mathbf{C}^3$中.

*§13.6　线性代数中的理想

在非交换环中，我们可以考虑"单边"理想. 环A中一个子集合L，如果x和y在L中，a在A中，则$x-y$与ax也在L中，那么称L是A的*左理想*. 右理想可以类似地定义. 与这些概念相对照，我们原来意义下的理想称为双边理想. 例如，在所有2×2矩阵组成的环M_2中，第一列都是零的所有矩阵构成左理想，但不能构成双边理想.

这些概念可以有效地应用到含有单位元素1的线性代数A中；正象§13.1中看到的，任意这样的线性代数是一个环. 在这种情况下，任意左理想L或右理想及关于数乘运算也是封闭的. 例如，如果ξ是L中任意元素，c是任意标量，那么L包含$c\xi$，这因为$c\xi=(c\cdot 1)\xi$是L中元素与A中某元素$c\cdot 1$的乘积. 如果把A看作它的标量域F上的线性空间，那么A的任意左(或右)理想是一个子空间.

如果一个线性代数没有真(双边)理想，则称它为*单代数*. 于是一个单代数没有真同态象.

定理 10　*域上所有$n\times n$矩阵组成的代数是单代数.*

证明　这个代数M_n以n^2个矩阵E_{ij}作为它的基底，E_{ij}是在(i,j)位置的元素为1，其余位置都是零. M_n中的一个真理想B将至少包含一个非零矩阵$A=\sum_{i,j}a_{ij}E_{ij}$(其中系数$a_{rs}\neq 0$). 那么每个矩阵

$$(a_{rs})^{-1}E_{kr}AE_{sk}=(a_{rs})^{-1}\sum_{i,j}E_{kr}E_{ij}E_{sk}a_{ij}=E_{kk} \qquad (26)$$

在B中．因此单位矩阵 $I=\sum_k E_{kk}$ 在B中，所以B一定是整个代数，于是B是假理想． 证毕

范德波恩（Wedderburn）（1908）证明了有名的定理 10 的逆定理．这个逆定理断言，特别复数域$\mathbf{C}$上的每个单代数与$\mathbf{C}$上所有$n\times n$矩阵组成的代数同构．为了讨论一般情形，我们需要可除代数的概念．可除代数是指这个线性代数是一个可除环．根据代数基本定理我们可以证明，复数域$\mathbf{C}$上唯一的可除代数是$\mathbf{C}$本身．著名的弗罗比尼乌斯定理断言，实数域$\mathbf{R}$上的可除代数只有$\mathbf{R}$，$\mathbf{C}$和四元数代数（§8.11）．

在任意可除环D上，我们可以构造一个任意n阶全矩阵代数$M_n(D)$，如下所述．可以按照普通法则

$$(a_{ij})+(b_{ij})=(a_{ij}+b_{ij}),$$

$$c(a_{ij})=(ca_{ij}), \tag{27}$$

$$(a_{ij})(b_{ij})=\left(\sum_{k=1}^{n}a_{ik}b_{kj}\right),$$

把系数在可除代数D中的两个$n\times n$矩阵相加和相乘．范德波恩的结果是，如果F是任意域，域F上最一般的单代数A可如下得到．取域F上任意可除代数D和任意正整数n，那么A是由系数在D中的所有$n\times n$矩阵组成．

习　题

1. 证明：每个可除代数是单代数．
2. 在可除代数中求出所有右理想．
3. 讨论一个环中左理想构成的代数，即描述它们的和，交与主左理想．
4. 证明：域F上线性代数的每个商环本身是一个线性代数．
5. (a) 证明：如果S是矢量空间F^n的子空间，以S中的矢量为行的

所有矩阵组成的集合是 $M_n(F)$ 的左理想.

*(b) 证明: $M_n(F)$ 的每个左理想 C 是 (a) 中所描述的一个理想. (提示: 证明 C 的矩阵的每一行是 C 中除了第一行外剩下的其他行都为零的矩阵的第一行. 利用 § 7.6 和 § 7.7 的方法.)

*6. 把定理 10 推广到任意可除环 D 上的全矩阵代数 $M_n(D)$.

§ 13.7 环的特征

任意环 R 可以看作加法群(阿贝耳群). 由任意 $a \in R$ 生成的循环子群是由 a 的 m 次幂组成, 其中 m 取整数. 用加法的记号, 我们把 a 的 m 次"幂"写成 $m \times a$. 于是, 如果 m 是正整数, 则

$$m \times a = a + a + \cdots + a \qquad (m\text{ 个被加项}). \tag{28}$$

如果 $m=0$, 则 $0 \times a = 0$, 而当 $m=-n$ 是负数时, 则有

$$\begin{aligned}(-n) \times a &= n \times (-a) \\ &= (-a) + (-a) + \cdots + (-a) \quad (n\text{ 个被加项}).\end{aligned} \tag{29}$$

我们称 $m \times a$ 是 a 的 m 项自然倍数, 它对任意 $m \in \mathbf{Z}$ 和任意 $a \in R$ 都有定义.

整环 D 中的元素的这些自然倍数具有任意交换环中幂的所有性质, 这些性质已经在 § 6.6 中证明过, 那里是按照乘法记号来叙述的. 因此有

$$\begin{aligned}(m \times a) + (n \times a) &= (m+n) \times a, \\ m \times (n \times a) &= (mn) \times a,\end{aligned} \tag{30}$$

以及

$$\begin{aligned}m \times (a+b) &= m \times a + m \times b, \\ m \times (-a) &= (-m) \times a.\end{aligned} \tag{31}$$

还有由分配律推出的一些性质. 其中一个一般分配律 (见 § 1.5) 是

$$(a+a+\cdots+a)b = ab+ab+\cdots+ab \qquad (m\text{ 个被加项}).$$

按照自然倍数表示这就是

$$(m\times a)b=m\times(ab)=a(m\times b). \tag{32}$$

当 $m=0$ 时，这个公式仍成立，当 m 是负数时公式也成立，这因为取 $m=-n$，定义(29)给出

$$(-n)\times ab=n\times(-ab)=[n\times(-a)]b=[(-n)\times a]b.$$

另一个一般分配律是法则

$$(a+\cdots+a)(b+\cdots+b)=ab+\cdots+ab.$$

它也可以改写成

$$(m\times a)(n\times b)=(mn)\times(ab). \tag{33}$$

对一切整数 m 和 n，不管正的，负的或零，这个公式都成立.

令 $a=1$ 是 R 的单位元素(乘法单位元素)，(32)表明 $m\times b$ 正是 $(m\times 1)b$，它是 b 与 1 的 m 次自然倍数的乘积. 此外，在(30)中令 $a=1$，我们看出，从 $\mathbf{Z}$ 到 R 的映射 $m\longmapsto m\times 1$ 保持和. 最后，在(33)中令 $a=b=1$，我们得到

$$(m\times 1)(n\times 1)=(mn)\times(1\cdot 1)=(mn)\times 1. \tag{33'}$$

这个映射保持积. 这就证明了

定理 11 对任意环 R，映射 $m\longmapsto m\times 1$ 是从 $\mathbf{Z}$ 到 R 的一个同态.

推论 1 任意环 R 中 1 的自然倍数组成的集合是一个同构于 $\mathbf{Z}$ 或 $\mathbf{Z}_m$ ($m>1$ 为某整数)的子环.

定义 环 R 的特征是指 R 的单位元素 1 的不同自然倍数 $m\times 1$ 的个数 m.

推论 2 在整环 D 的加法群中，所有非零元素具有相同的阶——即 D 的特征.

证明 对于所有非零元素 $b\in D$，$m\times b=0$ 当且仅当 $(m\times 1)b=0$，根据消去律，这等价于 $m\times 1=0$. 证毕

整数环 $\mathbf{Z}$ 具有特征①∞，而整数 $\mathbf{Z}_p$ 具有特征 p. ∞ 和 p 是唯

① 大多数作者用"特征 0"来代替"特征∞".

一可能的特征.

定理 12 一个整环的特征或者是∞或者是素数 p.

为证明定理用反证法. 我们假定某个整环 D 有有限特征, 这个特征是一个复合数 $m=rs$. 那么根据(33′), 环 D 的单位元素 1 满足

$$0=m\times 1=(rs)\times 1=(r\times 1)\cdot(s\times 1).$$

根据消去律, 或者 $r\times 1=0$, 或者 $s\times 1=0$. 因此 D 的特征一定是其中一个因子或者 r, 或者 s, 而不是我们所假定的 m.

推论 在任意整环中, 由单位元素生成的加法群是一个与 $\mathbf{Z}$ 或 $\mathbf{Z}_p$ 同构的子整环.

§1.5 的二项公式(9)说明了自然倍数的意义. 在任意交换环 R 中, 表达式

$$(a+b)^2=a^2+ab+ab+b^2=a^2+2\times(ab)+b^2$$

有一个中项, 确切地说它是自然倍数 $2\times(ab)$. 更一般地, 用归纳法可以证明 §1.5 中的二项公式(9), 它包含的二项系数是自然倍数, 于是我们可以写成

$$(a+b)^n$$
$$=a^n+\binom{n}{1}\times(a^{n-1}b)+\binom{n}{2}\times(a^{n-2}b^2)+\cdots+\binom{n}{n}\times b^n, \qquad (34)$$

其中系数 $\binom{n}{i}$ 是按公式

$$\binom{n}{i}=\frac{n!}{(n-i)!\,i!}, \qquad i=0,1,\cdots,n \qquad (35)$$

给出的自然数. 这里 $n!=n(n-1)\cdots 3\cdot 2\cdot 1$, 并且 $0!=1$.

定理 13 在素特征 p 的任意交换环 R 中, 对应 $a\longmapsto a^p$ 是一个同态.

证明 根据(6), 我们需要证明 $1^p=1$, 并且对所有 $a,b\in R$ 有 $(ab)^p=a^pb^p$, $(a\pm b)^p=a^p\pm b^p$. 前两个方程在任意交换环中都成

立. 为证明第三个方程, 在公式(34)和(35)中, 令 $n=p$. 因为 p 是素数, 所以它不能被 $i!$ 或 $(p-i)!$ $(0<i<p)$ 的任意因子整除. 因此(34)中适合 $0<i<p$ 的所有二项系数都是 p 的倍数. 但是环 R 具有特征 p, 因此 (34) 中所有含 $\binom{p}{i}$ $(0<i<p)$ 的项被去掉. 由此推出等式

$$(a\pm b)^p=a^p\pm b^p. \tag{36}$$

定理证毕.

推论 在特征为 p 的有限域 F 中, 对应 $a\longmapsto a^p$ 是一个自同构.

证明 因为在 F 中由 $a^p=0$ 可推出 $a=0$, 所以同态 $a\longmapsto a^p$ 的核是零, 而且这个同态是一对一的. 因为 F 是有限的, 所以这就意味着 $a\longmapsto a^p$ 还是映上的. 因此是一个自同构.

习　题

1. 证明: 对于正整数 m, 自然倍数 $m\times a$ 可以通过"递推公式" $1\times a=a$, $(m+1)\times a=m\times a+a$ 来定义.

2. 用归纳法证明关于正自然倍数的法则(30)和(32).

3. 作为定理 13 的一个推论来证明费马定理(§1.9 定理 18).

4. 关于有序整环的特征, 你能说些什么?

5. (a) 证明: 在特征为 p 的任意整环 D 中, $\alpha: a\mapsto a^p$ 是一一的(单一同态).

   (b) 证明: 如果 $D=\mathbf{Z}_p(x)$, 那么 α 的象是 D 的真子整环.

   (c) 证明: 有限域必有一个真自同构, 除非它是一个素域 $\mathbf{Z}_p$.

§13.8 域的特征

因为域被定义为除法(除零外)是可能的整环, 所以关于特征的讨论立刻可应用到域上. 如果域 F 的特征为 p, 那么根据定理 12, 由域 F 的单位元素生成的加法子群是一个子域, 并且它与由模

p 整数构成的有限域同构．如果域 F 的特征为 ∞，那么根据定理 12，由单位元素 1 生成的子群是由所有倍数 $m\times 1$ 组成，所以由 c 生成的子域是由所有商$\frac{m\times 1}{n\times 1}$(其中 $n\neq 0$) 组成．这个子域是所有倍数 $m\times 1$ 的子整环的商域．因此根据 § 2.2 的定理 7，它与有理数域同构，有理数域是整数 $m\longleftrightarrow m\times 1$ 构成的整环的商域．事实上，映射$\frac{m\times 1}{n\times 1}\longleftrightarrow\frac{m}{n}$是由 1 生成的子域与有理数域之间的同构．这就证明了下面的结果(参看 § 2.6 定理 18 的推论 2)：

定理 14 *在特征为∞的域中，由单位元素生成的子域与有理数域* **Q** *同构．*

同构$\frac{m\times 1}{n\times 1}\longleftrightarrow\frac{m}{n}$ 保持域 F 中的所有四种运算．于是在处理单个域 F 时，可以把每个商$\frac{m\times 1}{n\times 1}$与其相应的$\frac{m}{n}$等同起来．在这个约定之下，可以说每个特征为$\infty$的域包含着所有有理数$\frac{m}{n}$($n\neq 0$)．按照类似的约定，可以说每个特征为 p 的域包含域 $\mathbf{Z}_p$．在这种意义下，每个域都是极小域（即所谓素域）**Q** 和 $\mathbf{Z}_p$ 中的一个域的扩张．因此按照对已知域的不同扩张方式对域进行系统的分类是很自然的．这些将在下一章里讨论．

习　题

1．设 F_4 是恰有四个元素的任意域，

(a) 证明：F_4 的特征为 2．

(b) 证明：F_4 的不在素子域 $\mathbf{Z}_2$ 中的两个元素都满足 $x^2=x+1$．

(c) 用这个事实证明：F_4 与 § 13.3 中习题 8 所述的域 $\mathbf{Z}[\omega]/(2)$ 同构．

2．求出习题 1 的域 F_4 的全部自同构．

3．证明：关于二次方程的解的一般公式可以应用到特征为 2 的任意域上．

4．在什么样的域上，求解三次方程的一般公式(§ 5.5)仍然成立．

第十四章　代数数域

§14.1　代数扩张与超越扩张

剩下的两章讨论一般域F上的多项式方程$p(x)=0$的解及其性质．我们将证明任意这样的方程在F的适当的扩张中是可解的，这个扩张是指包含F作为子域的一个域K．例如$p(x)=0$在多项式环$F[x]$对于由p的倍式组成的主理想的商域$F[x]/(p)$中总有一个根．

在描述这种扩张的一般性质之后，我们将特别地研究有理数域$\mathbf{Q}$按这种方式扩张而得到的所有“代数数”构成的域．通过证明在某一个二次扩张$\mathbf{Q}[x]/(x^2-r)=q(\sqrt{r})$($r\in\mathbf{Z}$)中“整数”的唯一因子分解定理的问题，简短地介绍一下代数数论．例如，高斯整数$m+n\sqrt{-1}$($r=-1$的情形)可以唯一地分解成高斯素数．

域F最简单的一类扩张K是由单个元素$c\in K$的有理表达式$\dfrac{p(c)}{q(c)}=\dfrac{\sum a_kc^k}{\sum b_ic^i}$(系数$a_k$, $b_i\in F$)组成．例如，复数$a+bi$是由实数和单个复数i生成．未定元x的所有有理形式(具有有理系数)组成的域$\mathbf{Q}(x)$是由域$\mathbf{Q}$和元素x生成的．一个域可以按照几种不同方法生成．例如，域$\mathbf{Q}(\sqrt{2})$由方程$x^2-2=0$的根$\sqrt{2}$生成，它是由含有有理系数a和b的所有实数$a+b\sqrt{2}$组成（见§2.1的例子)．另一个不同的方程$x^2+4x+2=0$的一个根是$-2+\sqrt{2}$，它生成同一个域$\mathbf{Q}(\sqrt{2})$，因为这个域中的任意数可以按照这个新的生成元表示为

$$a+b\sqrt{2}=a+2b+b(-2+\sqrt{2}).$$

普通的配方方法应用到这个方程上得到 $x^2+4x+2=(x+2)^2-2$，所以 $y=x+2$ 满足新的方程 $y^2-2=0$，其根生成同一个域. 于是运用变量替换来化简方程对应着在相应的域中选取新的生成元.

让我们一般地描述由域 F 的任意扩张 K 中的已知元素生成的子域. 设 K 是给定的域，F 是 K 的子域，c 是 K 的一个元素. 考虑 K 中那些由形为

$$f(c)=a_0+a_1c+a_2c^2+\cdots+a_nc^n \quad (\text{每个 } a_i\in F) \qquad (1)$$

的多项式给出的元素. K 中包含 F 和 c 的任意子整环一定包含所有这样的元素 $f(c)$. 反过来，所有这样的多项式组成的集合在加法、减法和乘法运算之下是封闭的. 因此这些表达式(1)组成 K 中由 F 和 c 生成的子整环. 这个子整环一般都用带方括号的 $F[c]$ 来表示.

如果 $f(c)$ 和 $g(c)\neq0$ 是象(1)那样的多项式表达式，那么它们的商 $\dfrac{f(c)}{g(c)}$ 是 K 的元素，称为具有系数属于 F 的关于 c 的有理表达式. 所有这样的商组成的集合是一个子域；它是由 F 和 c 生成的域，一般都用带圆括号的 $F(c)$ 来表示.

如果域 K 是在它的子域 F 上由单个元素 c 生成的，则称 K 是 F 的**单扩张**，所以 $K=F(c)$. §2.1 所讨论的域 $\mathbf{Q}(\sqrt{2})$，$\mathbf{Q}(\sqrt[3]{5})$ 和 $\mathbf{Q}(\omega)$ 都是单扩张的例子. 可以证明，不管怎样，F 的任意扩张可以通过单扩张的一个有限或(良序)超限序列而得到.

在有理数域上，一些复数(如 i)，$\sqrt{2}$，$\sqrt[3]{5}$，$\sqrt{-3}$ 都满足有理系数多项式方程. 还有另外一些数，象 π 和 $e=2.71828\cdots$，可以证明它们不满足有理系数多项式方程(平凡情形除外). 后面这些数称为"超越数". 这种重要的分法可以应用到任意域的元素上.

定义 设 K 是任意域，F 是 K 的任意子域. 域 K 的元素 c，如果满足一个多项式方程

$$a_0+a_1c+a_2c^2+\cdots+a_nc^n=0 \quad (a_i\in F,\ 且不全为零).\ (2)$$

则称 c 在 F 上是代数的. K 的元素 c 如果在 F 上不是代数的，则称 c 在 F 上是超越的.

称单扩张 $K=F(c)$是 F 上的代数扩张还是超越扩张，根据生成元 c 在 F 上是代数的还是超越的而定. 单超越扩张的结构是特别容易描述的.

定理 1 如果 c 在 F 上是超越的，那么由 F 和 c 生成的子域 $F(c)$ 与系数在 F 中的关于未定元 x 的所有有理形式组成的域 $F(x)$ 同构. 同构可以选为 $a\longmapsto a$(每个 $a\in F$), $c\longmapsto x$.

证明 扩张 $F(c)$显然包含 F和系数在 F 中的所有有理表达式$\dfrac{f(c)}{g(c)}$. 如果 $F(c)$中的两个多项式表达式 $f_1(c)$和 $f_2(c)$相等，那么它们的系数一定逐项相等，因为如果不然，差 $f_1(c)-f_2(c)$将产生一个关于 c 的系数不全为零的多项式方程，这与 c 在 F 上是超越的假设相矛盾. 因此对应 $f(c)\longmapsto f(x)$是整环 $F[c]$ 和未定元 x 的多项式形式整环 $F[x]$之间的双射. 根据多项式运算法则，这个对应是一个同构. 根据 § 2.2 定理 6，它可以扩张成 $F(c)$ 和 $F(x)$ 之间的同构$\dfrac{f(c)}{g(c)}\longleftrightarrow\dfrac{f(x)}{g(x)}$.

习 题

1. 辨别下列各复数在有理数域 $\mathbf{Q}$ 上是代数的还是超越的，并给出证明：$\sqrt{7}$，$\sqrt[3]{5}$，π^2，$e+3$ (这里 $e=2.71828\cdots$)，$i+3$，$e^{2\pi i}$，$\sqrt{2}+i$.

2. 证明：如果 x 在 F 上是代数的，那么 x^2 和 $x+3$ 也是代数的，反之亦然.

3. $\mathbf{Q}(\sqrt{5})$中什么样的数生成整个域$\mathbf{Q}(\sqrt{5})$?

4. (a) 设 d 是非完全平方整数，描述域$\mathbf{Q}(\sqrt{d})$.

   (b) 求出$\mathbf{Q}(\sqrt{d})$中生成整个域的那些元素.

   (c) 把每个这样的元素表示为系数在 $\mathbf{Q}$ 中的二次方程的根.

§14.2 域上的代数元素

下面我们研究域F的单代数扩张的性质，该扩张是由F和在F上单个的代数元素u生成的．根据定义，这个元素必满足F上次数至少是1的多项式方程．同一个元素u可以满足很多不同的方程，例如，$\sqrt{2}$是$x^2-2=0$的根，也是$x^3-2x=0$, $x^4-4=0$等等方程的根．但是它恰好是一个首一不可约多项式方程的根（见下面习题6）．

定理2 如果域F的扩张K的一个元素u在F上是代数的，那么u是多项式整环$F[x]$中唯一的一个首一不可约多项式$p(x)$的零点．如果h是$F[x]$中另一个多项式，那么$h(u)=0$当且仅当在整环$F[x]$中，h是p的倍式，也就是当且仅当h在$F[x]$的主理想(p)中．

证明 满足$h(u)=0$的多项式$h\in F[x]$组成$F[x]$中的一个理想，这个理想恰好是"赋值映射"$p\longmapsto p(u)$所定义的同态ϕ_u: $F[x]\longrightarrow K$的核，这里映射$p\longmapsto p(u)$在$u\in K$处赋给每个多项式p一个值．象$F[x]$的所有理想一样，这个理想是主理想(§3.8定理11)，所以它由它的任意一个次数最低的元素的所有倍式组成．这些次数最低的元素中恰有一个是首一多项式，称它为p．这个p是不可约的，如果不然，它可以分解为$p=fg$，这里f和g是次数更小的多项式，由此推出$f(u)g(u)=p(u)=0$，所以或者$f(u)=0$，或者$g(u)=0$，这与选取p为适合$p(u)=0$的次数最低的多项式相矛盾．定理证完．

定义 在域F上代数的元素u的极小多项式是满足$p(u)=0$的(唯一的)首一不可约多项式$p\in F[x]$．u在F上的次数$n=[u:F]$是这个多项式的次数．

推论 如果元素u在F上的次数为n，那么我们有，a_0+a_1u

$+\cdots+a_{n-1}u^{n-1}=0(a_i\in F)$ 当且仅当 $a_0=a_1=\cdots=a_{n-1}=0$.

现在我们有可能描述 K 的由 F 和上述的代数元素 u 所生成的子域. 这个子域 $F(u)$ 显然包含由所有可表为系数在 F 中的多项式 $f(u)$ 的元素组成的子整环 $F[u]$, 此外还将证明, 映射 $f(x)\longmapsto f(u)$ 是商环 $F[x]/(p)$ 与 $F(u)$ 之间域的同构 ϕ': $F[x]/(p)\longrightarrow F(u)$.

这一节余下的部分将讨论这个结论. 由多项式的加法和乘法公式, 显然有 ϕ' 是由 $F[x]$ 到子整环 $F[u]$ 的一个满同态. 但是实际上, 整环 $F[u]$ 是一个子域. 事实上, 让我们求 $F[u]$ 中任意元素 $f(u)\neq 0$ 的逆. 不等式 $f(u)\neq 0$ 意味着 u 不是 $f(x)$ 的根, 因此根据定理 2, $f(x)$ 不是不可约多项式 $p(x)$ 的倍式, 所以 $f(x)$ 与 $p(x)$ 互素. 因此我们可以写

$$1=t(x)f(x)+s(x)p(x), \tag{3}$$

这里 $t(x)$ 和 $s(x)$ 是 $F[x]$ 中适当的多项式. $F[u]$ 中相应的方程是 $1=t(u)f(u)$. 这就表明, $F[u]$ 的非零元素 $f(u)$ 有一个逆元素 $t(u)$, $t(u)$ 也是 u 的多项式①, 于是就证明了 $F[u]$ 是 K 的子域.

反之, 因为 K 的每个包含 F 和 u 的子域显然包含 $F[u]$ 中的每个多项式 $f(u)$, 所以我们看出 $F[u]$ 是 K 的由 F 和 u 生成的子域. 我们证明了

定理 3 设 K 是任意域, u 是 K 的一个元素, 它在 K 的子域 F 上是代数的; 设 $p(x)$ 是以 u 为其根的 F 上首一不可约多项式. 那么从多项式整环 $F[x]$ 到 $F(u)$ 的映射 ϕ': $f(x)\longmapsto f(u)$ 是以 $(p(x))$ 为核的满同态.

把这个定理同 §13.3 定理 5 的推论 2 结合起来, 我们有一个

---

① 例如, 在 $Q(\sqrt{3})$ 中, $1+\sqrt{3}$ 有乘法逆, 通过分母有理化可以求得这个逆是

$$\frac{1}{1+\sqrt{3}}=\frac{1-\sqrt{3}}{(1+\sqrt{3})(1-\sqrt{3})}=-\frac{1}{2}+\frac{1}{2}\sqrt{3}.$$

直接推论.

定理 4 在定理 3 中, $F(u)$ 与商环 $F[x]/(p)$ 同构, 这里 p 是 u 所满足的域 F 上首一不可约多项式.

商环 $F[x]/(p)$ 可以描述得非常简单. 每个多项式 $f(x)\in F[x]$ 在模 (p) 之下与它用 $p(x)$ 除所得的余式 $r(x)=f(x)-a(x)p(x)$同余, 这个余式是次数小于 n 的唯一的多项式

$$r(x)=r_0+r_1x+\cdots+r_{n-1}x^{n-1}. \tag{4}$$

把两个这样的多项式相加或相减, 恰好是对它们的相应的系数相加或相减. 为把它们相乘, 先按照 § 3.1 的(3′) 计算出多项式乘积, 然后再计算用 $p(x)$除时所得的余式.

例如, 有理数域 $F=\mathbf{Q}$ 通过 $u=\sqrt{2}$扩张成 $\mathbf{Q}(\sqrt{2})$, 在这个特殊情形中, 我们有 $p(x)=x^2-2$. 因此 $\mathbf{Q}(u)$ 的任意元素可以写成$a+b\sqrt{2}$, 其中 a 和 b 为有理数, 并且

$$\begin{aligned}&(a+b\sqrt{2})(c+d\sqrt{2})\\&=ac+(ad+bc)\sqrt{2}+bd(\sqrt{2})^2\\&=(ac+2bd)+(ad+bc)\sqrt{2}.\end{aligned}$$

公式(4)表明商环 $F[x]/(p)$是 F 上一个 n 维矢量空间, 它是有限维矢量空间 $F[x]$ 对由 $p(x)$ 的倍式组成的子空间的商空间. 还要注意, 乘法是双线性的(对每个因子是线性的). 因此代数扩张 $F[x]/(p)$ 在 § 13.1 的意义下也可以看作F 上的交换线性代数.

习 题

1. 求出$\sqrt{3}$所满足的五个不同多项式方程, 并证明它们都是$\sqrt{3}$ 所满足的首一不可约多项式(域 $\mathbf{Q}$ 上)的倍式.

2. 在由不可约方程 $u^3-6u^2+9u+3=0$ 的根 u 生成的单扩张 $\mathbf{Q}(u)$中, 按照元素 $1, u, u^2$ 把下列各元素表示成(4)中的形式:

$$u^4,\ u^5,\ 3u^5-u^4+2,\ \frac{1}{u+1},\ \frac{1}{u^2-6u+8}.$$

3. 在由 $x^3+2x+2=0$ 的根 u 生成的单扩张 $\mathbf{Q}(u)$ 中，把下列各元素表示成(4)中的形式：$(u^3-2)(u^3+3u)$，$u^4(u^4+3u^2+7u+5)$，$\frac{1}{u}$，$\frac{u+2}{u^2+3}$.

4. 把复数域表示为由所有实系数多项式组成的整环 $\mathbf{R}[x]$ 得到的商环.

5. 把域 $\mathbf{Q}(\sqrt{2})$ 表示为由有理系数多项式组成的整环 $\mathbf{Q}[x]$ 得到的商环.

6. 根据有关的定义直接证明：如果 u 在 F 上是代数的，那么以 u 为根的次数最低的首一多项式在 F 上是不可约的.

7. 根据有关的定义证明：如果 u 是域 K 的任意元素，F 是 K 的任意子域，那么以 u 为根的，系数在 F 中的所有多项式 $g(x)$ 的集合是 $F[x]$ 的一个理想.

§14.3 根的添加

迄今，我们假定已经给出域 F 的扩张 K，并描述了 K 的由 F 和已知元素 $u\in K$ 生成的子域，这里 u 是由 F 上的极小多项式（即首一不可约多项式）p 给出的，它满足 $p(u)=0$. 另一方面，我们恰恰可以从 F 和不可约多项式 p 出发，构造一个包含 $p(x)=0$ 的根的较大的域. 这种构造方法是把第五章用过的由实数域 $\mathbf{R}$ 添加上方程 $x^2+1=0$ 的一个虚根来构造复数域 $\mathbf{C}$ 的方法加以一般化. 定理 3 和定理 4 指出在一般情形下怎样得到同样的结果.

定理 5 *如果 F 是域，p 是 F 上不可约多项式，那么存在域 $K\cong F[x]/(p)$，它是由 $p(x)$ 的根 u 生成的 F 的单代数扩张.*

证明 因为 $p(x)$ 是不可约的，所以主理想 (p) 是 $F[x]$ 中的极大理想. 因此根据 §13.3 定理 6，商环 $F[x]/(p)$ 是一个域. 它包含 F 和剩余类 $x+(p)$（该剩余类包含 x），并满足在 $F[x]/(p)$ 中有 $p(x)=0$.

这个单扩张除同构外是唯一的.

定理 6 如果域 $F(u)$ 和 $F(v)$ 是同一域 F 的两个单代数扩张，它们分别由 F 上同一个不可约多项式 p 的根 u 和 v 生成，那么 $F(u)$ 与 $F(v)$ 同构。特别恰好存在一个 $F(u)$ 到 $F(v)$ 的同构，在这个同构之下，u 对应于 v，F 的每个元素与自身对应。

证明 由定理 4 提供的同构

$$F(u)\xleftarrow{\phi_u}F[x]/(p)\xrightarrow{\phi_v}F(v)$$

取它们的合成 $\phi_u^{-1}\phi_v$.

定理 5 可以用来构造各种有限域。例如，从模 3 整数的域 $\mathbf{Z}_3$ 出发，对于多项式 x^2-x-1，0，1，2 三个元素没有一个是它的零点，因此它在 $\mathbf{Z}_3[x]$ 中是不可约的。所以商环 $\mathbf{Z}_3[x]/(x^2-x-1)$ 是一个域 K，它是由它的子域 $\mathbf{Z}_3$ 和 x 的陪集（称为 u）生成的。而且因为 $[u:F]=2$，所以这个域 K 的每个元素可以唯一地写成 $a+bu$，其中 $a,b\in F$，因此 K 恰好有九个元素。

这个域还可以不用商环概念直接来构造。它刚好由九个形为 $a+bu$ 的元素组成。它们之中两个元素之和由法则

$$(a+bu)+(c+du)=(a+c)+(b+d)u$$

给出。为计算两个这种类型的元素之积，我们可先按自然方式乘出来，然后再根据已给出的方程 $u^2=u+1$ 来化简。其结果是

$$\begin{aligned}(a+bu)(c+du)&=ac+(ad+bc)u+bdu^2\\&=(ac+bd)+(ad+bc+bd)u.\end{aligned}$$

我们可以详细验证，这九个元素 $a+bu(a,b\in\mathbf{Z}_3)$ 在上述两种运算之下满足域的所有公设。特别是，非零元素的逆由下表给出

1	2	u	$2u$	$1+u$	$1+2u$	$2+u$	$2+2u$
1	2	$2+u$	$1+2u$	$2+2u$	$2u$	u	$1+u$

根据上述构造，这个域显然是由剩余类域 $\mathbf{Z}_3$ 添加 u 生成的域

$\mathbf{Z}_3(u)$. 它是有限域中最简单的例子之一(见§15.3).

上述添加方式可以用到任意基域F上. 如果F是实数域$\mathbf{R}$, $p(x)$是$\mathbf{R}$上不可约多项式x^2+1,那么这个构造得到域$\mathbf{R}(u)$, 它是由满足$u^2=-1$的数u生成. 这个数u的性质很象$i=\sqrt{-1}$, 并且域$\mathbf{R}(u)$实际上与复数域$\mathbf{C}$同构, 这同我们在第五章中用过的从实数域得到复数域的构造方法稍微有些不同.

如果F是模p整数的域$\mathbf{Z}_p$, $p(x)$是F上某一不可约多项式, 则上面的构造方法将产生一个由元素$a_0+a_1u+\cdots+a_{n-1}u^{n-1}$组成的域. 因为每个系数$a_i$只有$p$种(有限)选择, 因此这样构造出的域是具有$p^n$个元素的有限域, 这里$n$是多项式$p(x)$的次数.

用同样的方法, 我们还可以构造代数函数域. 例如, 设$F=\mathbf{C}(z)$是所有有理复函数组成的域, 假设我们要求把满足$t^2=(z^2-1)(z^2-4)$的函数$t(z)$添加到F上. 我们可以把多项式$p(t)=f(z,t)=t^2-(z^2-1)(z^2-4)$看作系数在$\mathbf{C}(z)$中的$t$的二次不可约多项式. 那么商环$K=F[t]/(p(t))$是一个包含所有有理函数和代数函数$t$的域. 我们可以把$t(z)$作为$K$的一个元素来研究,而不必对它(它是双值的)构造黎曼面. 域K称为椭圆函数域,因为它是由椭圆积分

$$\int\sqrt{(z^2-1)(z^2-4)}\,dz$$

的被积函数生成的.

如果把定理6应用到象x^3-5这样的普通多项式(它在有理数域$\mathbf{Q}$上不可约)上, 它可以得到由正的$\sqrt[3]{5}$生成的$\mathbf{Q}$的扩张$\mathbf{Q}(\sqrt[3]{5})$, 也可得到扩张$\mathbf{Q}(\omega\sqrt[3]{5})$,这里$\omega=\dfrac{-1+i\sqrt{3}}{2}$是复三次单位根. 可以证明这两个域$\mathbf{Q}(\sqrt[3]{5})$和$\mathbf{Q}(\omega\sqrt[3]{5})$在代数上没有什么区别,因为它们是同构的.

粗略地说，这个同构意味着一个不可约多项式 $p(x)$ 的任意两个根具有相同的性质，根 u 的所有代数性质都可以从它所满足的不可约方程推导出来. 有很多这样的同构例子. 例如，复数域 $\mathbf{C}=\mathbf{R}(i)$ 是在实数域 $\mathbf{R}$ 上添加方程 $x^2+1=0$ 的两个根 $\pm i$ 中任意一个而生成的，因此根据定理 6，存在一个把 i 映射到 $-i$ 的 $\mathbf{C}$ 的自同构. 这个自同构刚好是一个数和它的通常共轭复数之间的对应 $a+bi\longleftrightarrow a-bi$.

习　题

1. 列出下列各域的非恒等对应的自同构：$\mathbf{Q}(\sqrt{2})$，$\mathbf{Q}(\sqrt{-3})$，$\mathbf{Q}(i)$.

2. 分别列出一个由复数组成的非实域与实域 $\mathbf{Q}(\sqrt[3]{5})$ 和 $\mathbf{Q}(\sqrt[4]{2})$ 同构.

3. 证明：x^3+x-1 在模 5 整数域 $\mathbf{Z}_5$ 上是不可约的. 如果把这个多项式的根添加到 $\mathbf{Z}_5$ 上，那么所得到的域有多少个元素？

4. (a) [①]求出在模 2 整数域 $\mathbf{Z}_2$ 上不可约的 2 次和 3 次多项式.

   (b) 对四元素域构成加法表和乘法表.

5. (a) 证明：正文中构造的九元素域的特征是 3.

   (b) 对这个域明显地列出同构 $a\longleftrightarrow a^3$.

6. (a) 求出域 $\mathbf{Z}_3$ 上所有二次不可约多项式.

   (b) 证明：任意两个九元素域同构.（提示：首先证明这样的有限域中每个元素在 $\mathbf{Z}_3$ 上是二次的.）

7. 证明：t 的多项式 $t^2-(x^2-1)(x^2-4)$ 在域 $\mathbf{C}(x)$ 上是不可约的.（提示：用 § 3.9 的结果.）

8. 证明：正文中的椭圆函数域 $\mathbf{C}(x,y)$ 可以通过在 $\mathbf{C}(x)$ 上添加方程 $t^2=\dfrac{x^2-4}{x^2-1}$ 的根而得到.

9. 如果 $g(t)$ 是可约多项式，那么在商环 $F[t]/(g(t))$ 中哪些元素确有逆元素？

10. 用 § 13.3 定理 6 给出关于 $F[t]/(p(t))$ 是域的另一个证明.

---

① 原书此题有误，现按第三版译出. ——译者注.

§14.4 次数与有限扩张

在一个 n 次元素 u 生成的单扩张 $F(u)$中，每个元素 w 按公式(4)有唯一的表达式为

$$w=a_0+a_1u+\cdots+a_{n-1}u^{n-1}, \tag{5}$$

其中系数在 F 中. 这唯一的表达式同一个矢量按照基矢量 1, $u, \cdots, u^{n-1}$ 的表达式极为相似. 这就暗示我们运用矢量空间的概念. 的确，域 F 的任意扩张可以看作域 F 上的矢量空间：只要不管域 K 的元素的乘法，而把 K 的两个元素相加和 K 的元素同 F 的元素的"数乘"两种运算当作矢量空间的运算. 这些加法和数乘运算满足矢量空间的所有公设. 如果这个矢量空间 K 是有限维的，那么称域 K 是 F 的*有限扩张*，并且把这个矢量空间的维数 n 称为*扩张次数* $n=[K:F]$.

例如，复数域 $\mathbf{C}=\mathbf{R}(i)$是实子域 $\mathbf{R}$(象在§5.2中那样) 上的二维矢量空间；由有理数域 $\mathbf{Q}$ 和 5 的三次根生成的域 $\mathbf{Q}(\sqrt[3]{5})$ 是有理子域 $\mathbf{Q}$ 上的三维矢量空间，等等. 一般地，关于单代数扩张的定理 4 可以按照维数重述如下.

定理 7 *域 F 上代数元素 u 的次数，等于把扩张 $F(u)$ 看作 F 上矢量空间时 $F(u)$的维数. 这个矢量空间有一组基底 1, $u, \cdots, u^{n-1}$.*

在§14.5 中，我们将要说明如何用矢量空间的方法来分析由域 F 添加几个不同代数元素而得到的扩张. 但是在讨论这样的"多重扩张"之前，我们首先来看一下，这种矢量空间的方法怎样能使我们比较 F 的同一个单代数扩张 $F(u)$ 中的不同元素所满足的不可约方程.

关于矢量空间的一个基本事实是维数的不变性 (矢量空间的任意两组基元素的个数相同). 这个事实可以应用到域的有限扩

张这种特殊情形，如下所述.

推论 *如果域F上的两个代数元素u和v生成同一个扩张$F(u)=F(v)$，那么u和v在F上的次数相同.*

一个单代数扩张是有限扩张，反之，每个有限扩张是由代数元素组成的.

定理 8 *F的有限扩张K的每个元素w在F上是代数的，并且满足一个次数至多是n的F上不可约方程，这里$n=[K:F]$是给定的扩张的次数.*

证明 给定元素w的$n+1$个幂$1, w, w^2, \cdots, w^n$是n维矢量空间K的元素，因此在F上一定线性相关（§7.4 定理5的推论2）. 所以必有线性关系$b_0+b_1w+\cdots+b_nw^n=0$，其中系数不全为零. 可以把它解释为多项式，于是这个关系就意味着w在F上是代数的.

推论 *单代数扩张$F(u)$的每个元素在F上是代数的.*

这个重要的结论使我们确信，超越元素决不能出现在一个单代数扩张中.

在讨论一个特殊的单代数扩张$F(u)$时，要系统地应用u所满足的不可约多项式$p(x)$. 根据定理2，这个扩张中的元素$g(u)$是零当且仅当多项式$g(x)$可被$p(x)$整除. 例如，假定$\mathbf{Q}(u)$是有理数域$\mathbf{Q}$上的三次扩张，它是由x^3-2x+2的一个根u生成. 根据爱森斯坦不可约准则（§3.10），这个多项式是不可约的. 这个扩张$\mathbf{Q}(u)$中的元素$w=u^2-u$一定满足某个次数至多是3的多项式方程. 为求出这个方程，象在定理4中那样，按照$1, u$和u^2把幂$w^2=u^4-2u^3+u^2$，$w^3=u^6-3u^5+3u^4-u^3$线性地表示出来. 反复运用已知的方程$u^3=2u-2$，这是可以做到的. 由此得到

$$w=u^2-u,\ w^2=3u^2-6u+4,\ w^3=16u^2-28u+18.$$

$1, w, w^2$和w^3之间一定满足一个线性关系，为得到这个关系，我们

可以由前两个线性方程解出 u 和 u^2 为

$$
\begin{aligned}
u&=-\frac{w^2}{3}+w+\frac{4}{3},\\
u^2&=-\frac{w^2}{3}+2w+\frac{4}{3}.
\end{aligned}
\tag{6}
$$

把这些代入 w^3 的表达式中就得到所需要的方程

$$w^3-4w^2-4w-2=0.$$

根据爱森斯坦定理，这个方程是 $\mathbf{Q}$ 上不可约的，根据方程(6)我们也可以说 u 在 $\mathbf{Q}(w)$ 中，所以 $\mathbf{Q}(u)=\mathbf{Q}(w)$，于是 u 和 w 生成同一个扩张，根据定理 7 的推论可知，它们在 $\mathbf{Q}$ 上的次数都是 3．这就意味着 w 所满足的任意三次方程一定是不可约的．

习　题

1．下列每个数都在 $\mathbf{Q}$ 的一个单代数扩张中，因此在 $\mathbf{Q}$ 上是代数的．对每种情形，求出该数所满足的首一不可约方程．

(a) $2+\sqrt{3}$，(b) $\sqrt[4]{5}+\sqrt{5}$，(c) $\sqrt[3]{2}+\sqrt[3]{4}$，

(d) u^2-1，这里 u 满足 $u^3=2u+2$，

(e) u^2+u，这里 u 满足 $u^3=-3u^2+3$．

2．证明：实数域 $\mathbf{R}$ 的每个有限扩张或者是 $\mathbf{R}$ 本身，或者同构于复数域 $\mathbf{C}$．

3．证明：复数域没有真有限扩张．

4．(a) 证明：如果 K 是有理数域 $\mathbf{Q}$ 的二次扩张，那么 $K=\mathbf{Q}(\sqrt{d})$，其中 d 是一个整数，非完全平方且无平方因子．

(b) 如果域 $\mathbf{Q}$ 用特征为 ∞ 的域 F 代替，那么上述结果还正确吗？如果用任意特征的域代替，情况如何？

5．未定元 x 的有理形式构成的域 $F(x)$ 是 F 的有限扩张吗？为什么？

6．证明：特征为 p 的有限域中元素的个数是 p 的幂．

7．(a) 证明：在模 p 整数域 $\mathbf{Z}_p$ 上恰有 $\dfrac{p^2-p}{2}$ 个二次首一不可约多项式．

(b) 证明：对每个 p，都存在一个含有 p^2 个元素特征为 p 的域.

*8. 证明：在模 p 整数域 $\mathbf{Z}_p$ 上恰有 $\frac{p^2-p}{3}$ 个三次首一不可约多项式.

*9. 设 F 是包含在整环 D 中的一个域，证明：

(a) D 是 F 上的矢量空间.

(b) 如果把 D 看作 F 上的矢量空间，它是有限维的，那么 D 是一个域.

§14.5 多重代数扩张

域 F 的有限扩张可以通过一系列的单扩张来构造. 如果 F 的特征是 ∞，那么我们可以证明，任意这样的多重扩张可以表示成一个单扩张，也就是说，它是在 F 上添加一个适当选择的单个元素而生成的. 我们将略去此证明，而直接来讨论多重扩张的性质. 一般地，如果 K 是 F 的包含元素 $c_1, c_2, \cdots, c_r$ 的任意扩张，那么记号 $F(c_1, c_2, \cdots, c_r)$ 表示由 $c_1, \cdots, c_r$ 和 F 的元素生成的 K 的子域（这个子域是由系数在 F 中关于 $c_1, \cdots, c_r$ 的有理形式所表示的所有元素组成的）. 另一方面，这样的多重扩张可以反复进行单扩张而得到. 例如，$F(c_1, c_2)$ 是单扩张 $L=F(c_1)$ 的单扩张 $L(c_2)$.

在求解方程时可以产生多重代数扩张，在这里引进适当的辅助方程常常是有用的. 例如，方程 $x^4-2x^2+9=0$ 可以写成

$$x^4-2x^2+9=(x^4-6x^2+9)+4x^2$$
$$=(x^2-3)^2+4x^2=0.$$

所以这个方程变为 $\left(\frac{x^2-3}{2x}\right)^2=-1$. 这个公式表明，包含上述给定方程的根 u 的任意域，也包含方程 $y^2=-1$ 的根 $i=\frac{u^2-3}{2u}$. 如果我们把辅助量 i 添加到有理数域 $\mathbf{Q}$ 上，那么原来的方程在 $\mathbf{Q}(i)$ 上就变为可约的，因为

$$x^4-2x^2+9=(x^2-3+2xi)(x^2-3-2xi).$$

根据普通的公式，因式 $x^2-3-2ix$ 有一个根 $u=i+\sqrt{2}$. 于是

原来方程在域 $K=\mathbf{Q}(\sqrt{2},i)$ 中就有一个根 u. 这个域 K 可以在 $\mathbf{Q}$ 上先添加$\sqrt{2}$,后添加 i 而得到. 中间域 $\mathbf{Q}(\sqrt{2})$是由实数组成的, 因此不可能包含 i. 所以 i 所满足的二次方程 $y^2+1=0$ 在实域 $\mathbf{Q}(\sqrt{2})$上一定仍然是不可约的,所以扩张 $\mathbf{Q}(\sqrt{2},i)$ 在 $\mathbf{Q}(\sqrt{2})$上的次数是 2,它的两个基元素是 1 和 i.而域 $\mathbf{Q}(\sqrt{2})$ 在 $\mathbf{Q}$ 上有一组基底 $1,\sqrt{2}$. 所以在整个域 $\mathbf{Q}(\sqrt{2},i)$ 中的任意元素 w 可以表示成

$$
\begin{aligned}
w&=(a+b\sqrt{2})+(c+d\sqrt{2})i\\
&=a+b\sqrt{2}+ci+d\sqrt{2}i,
\end{aligned}
\tag{7}
$$

其中 a,b,c,d 是有理数. 于是 $1,\sqrt{2},i,\sqrt{2}i$ 这四个元素构成 $\mathbf{Q}$ 上整个扩张 $K=\mathbf{Q}(\sqrt{2},i)$的一组基底. 这种计算基底的方法可以一般地叙述如下:

定理 9 如果元素 $u_1,\cdots,u_n$ 构成 F 的有限扩张 K 的一组基底, 而 $w_1,\cdots,w_m$ 组成 K 的扩张 L 的一组基底, 那么 mn 个乘积 $u_iw_j(i=1,\cdots,n;j=1,\cdots,m)$ 构成 F 的扩张 L 的一组基底.

证明 L 中任意元素 y 可以表示成给定基底的线性组合 $y=\sum_j r_jw_j$, 这里系数 $r_j\in K$. 每个系数 r_j 又可以表示成 K 的这组基元素的某线性组合 $r_j=\sum_i a_{ij}u_i$, 这里每个系数 $a_{ij}\in F$. 代入这些值,得到

$$
y=\sum_j\sum_i a_{ij}u_iw_j,
$$

这表现为已假定的元素 u_iw_j 的一个线性组合, 这里系数在 F 中. 用同样类型的逐次论证方法可以证明, 这 mn 个元素在 F 上是线性无关的,因此它们构成 L 的一组基底. 证毕

由定理 9 可以得出很多推论. 首先, 我们可以把与所用的特

殊基底无关的结果叙述如下：

推论 1 如果K是F的有限扩张，L是K的有限扩张，那么L是F的有限扩张，它的次数是

$$[L:F]=[L:K][K:F] \quad (L\supset K\supset F). \tag{8}$$

推论 2 如果K是F的次数为$n=[K:F]$的有限扩张，那么K的每个元素u在F上的次数是n的因子.

证明 元素u生成单扩张$F(u)$，因此根据(8)式有$n=[K:F(u)][F(u):F]$，这里第二个因子是我们所考虑的元素u的次数.

推论 3 有限扩张$K\supset F$的元素u生成整个扩张当且仅当$[K:F]=[u:F]$.

证明 如果u在F上满足一个次数为$[K:F]$的不可约方程，那么u在F上生成n次子域$F(u)$. 根据(8)式，这个子域一定包含整个K.

推论 4 如果$K=F(y_1,y_2,\cdots,y_r)$是由r个量$y_1,\cdots,y_r$生成的域，这里逐个y_i在由前$i-1$个量生成的域$F(y_1,\cdots,y_{i-1})$上是代数的，那么K是F的有限扩张，K中的每个元素在F上是代数的.

证明 每个次数$[F(y_1,\cdots,y_{i-1},y_i):F(y_1,\cdots,y_{i-1})]$是有限的，因此根据推论 1，整个次数$[K:F]$是有限的. 根据定理 8，$K$中每个元素在$F$上是代数的.

推论 5 如果$p(x)$是域F上一个三次不可约多项式，K是F的2^m次扩张，那么$p(x)$在K上是不可约的.

这个推论特别意味着，三次不可约方程决不能通过逐次求平方根的方法来解，这是因为，把一个平方根添加到域F上，或者全然没有给出扩张，或者给出二次扩张，所以由任意多个平方根得到的扩张$K=F(\sqrt{a},\sqrt{b},\sqrt{c},\cdots)$的次数是2的某个幂$2^m$. 根

据推论 5，这个扩张决不包含给定三次不可约方程的根.

为证明推论，假定 $p(x)$ 在 2^m 次的域 K 上是可约的. 那么三次多项式 $p(x)$ 一定至少有一个线性因子 $x-u$，于是 K 包含 $p(x)$ 的根 u. 但是根据推论 2，这种在 F 上的次数为 3 的元素 u 不可能包含在 F 上的次数为 2^m 的域 K 中. 这就证明了 $p(x)$ 在 K 上是不可约的.

这个推论是下述定理的代数基础：只用直尺和圆规不可能解一般的倍立方或三等分任意角这样的经典问题. 任意这样的作图问题可以化为解析的形式. 这个问题的对象是由一些点和直线组成. 对于某一组坐标轴，这些点的坐标（和这些直线的方程中系数之比）是一个实数集合，它生成某个由实数组成的域 F. 在用直尺和圆规作图时，每一步都提供某些新的点和直线. 可以证明[①]，相应的新的数域或者是 F 本身或者是 F 的二次扩张，因此重复上述作图便产生点和直线的集合，它对应于 F 上的 2^m 次的域 K.

现在考虑倍立方问题. 问题的对象是由三个坐标轴，每个轴上的一个单位线段，以这些线段为边的一个立方体组成. 这个问题是作另外一个具有两倍体积的立方体. 这个新立方体的边长将满足方程 $x^3-2=0$. 根据爱森斯坦定理，这个方程在有理数域 $\mathbf{Q}$ 上是不可约的（这个域 $\mathbf{Q}$ 与问题的对象有联系）. 根据推论 5，在对应于直尺和圆规作图的域 K 上，多项式 x^3-2 还是不可约的. 因此通过这种方法不可能构造出（比如说沿 x 轴）一个线段，使它是倍立方体的边.

三分角问题可按类似的方法处理，问题的实质在于，写出一个用整个角的余弦来表示三分之一角的余弦的三角方程. 对于大多数角来说，这又给出一个三次不可约方程.

---

① 本质上，这依赖于下述事实：圆（圆规）的方程是二次的，直线（直尺）的方程是线性的.

习　题

1. 在定理 9 中，详细证明：mn 个元素 $u_i w_j$ 在 F 上线性无关.

2. 证明：正文中所处理的方程 x^4-2x^2+9 在 $\mathbf{Q}$ 上是不可约的.（提示：用域 $\mathbf{Q}(\sqrt{2},i)$ 的次数证明.）

3. 证明：如果 $p(x)$ 是 F 上 q 次不可约多项式，K 是 F 的有限扩张，其次数与 q 互素，那么 $p(x)$ 在 K 上是不可约的.

4. 确定有理数域 $\mathbf{Q}$ 上的下列各多重扩张的次数，并说明理由.

　(a) $\mathbf{Q}(\sqrt{3},i)$，　(b) $\mathbf{Q}(\sqrt[3]{5},\sqrt{-2})$，

　(c) $\mathbf{Q}(\sqrt{18},\sqrt[4]{2})$，　(d) $\mathbf{Q}(\sqrt{8},3+\sqrt{50})$，

　(e) $\mathbf{Q}(\sqrt[3]{2},u)$，这里 u 满足 $u^4+6u+2=0$，

　(f) $\mathbf{Q}(\sqrt{3},\sqrt{-5},\sqrt{7})$，　(g) $\mathbf{Q}(\sqrt{3},\sqrt{2})$.

5. 对习题 4 中的每个域给出 $\mathbf{Q}$ 上的一组基底.

6. 确定下列多项式在指定的域上是否是不可约的，并给出理由.

　(a) x^2+3，在 $\mathbf{Q}(\sqrt{7})$ 上；

　(b) x^2+1，在 $\mathbf{Q}(\sqrt{-2})$ 上；

　(c) x^3+8x-2，在 $\mathbf{Q}(\sqrt{2})$ 上；

　(d) x^5+3x^3-9x-6，在 $\mathbf{Q}(\sqrt{7},\sqrt{5},1+i)$ 上.

7. 在下列每种情形中，确定给出的数 u 是否生成给出的有理数域 $\mathbf{Q}$ 的扩张. 对每种情形证明你的答案是正确的.

　(a) $u=\sqrt[3]{7}$，在 $\mathbf{Q}(\sqrt[3]{7})$ 中；

　(b) $u=\sqrt{2}+\sqrt{5}$，在 $\mathbf{Q}(\sqrt{2},\sqrt{5})$ 中；

　(c) $u=2+\sqrt[3]{9}$，在 $\mathbf{Q}(\sqrt[3]{3})$ 中；

　(d) $u=\dfrac{\sqrt{2}-1}{1+\sqrt{2}}$，在 $\mathbf{Q}(\sqrt{2})$ 中；

　(e) $u=v^2+v+1$，在 $\mathbf{Q}(v)$ 中，这里 v 满足 $v^3+5v-5=0$.

8. $c=\pi^6+5\pi^3+2\pi-14$ 在有理数域 $\mathbf{Q}$ 上是超越的还是代数的？为什么？

9. 证明：如果 K 是 F 的素数次扩张，那么不在 F 中的 K 中任意元素在 F 上生成整个 K.

10. (a) 求用 $\cos 3\theta$ 给出 $\cos\theta$ 的三次方程.

(b) 证明：当 $3\theta=60^\circ$ 时，这个方程在 $\mathbf{Q}$ 上是不可约的（这意味着 60° 角不能用直尺和圆规三等分）。

§14.6 代 数 数

代数数 u 是满足含有不全为零有理系数的多项式方程的复数. 即

$$a_0+a_1u+a_2u^2+\cdots+a_nu^n=0,\quad (a_i\in\mathbf{Q},\ a_i\ \text{不全为零}).\qquad (9)$$

换句话说，代数数是在有理数域 $\mathbf{Q}$ 上代数的任意复数. 在讨论域的扩张时，我们反复用到代数数的例子，例如 i，$\sqrt{-2}$，$\sqrt[5]{3}$，或 ω.

定理 10 *所有代数数组成的集合是可数的.*

验证这个命题需要我们描述一下对所有代数数进行计数或排列的方法. 首先，我们把它们所满足的方程都列出来. 我们注意，代数数所满足的方程(9)可以用它的有理系数的公分母去乘，于是得到一个含有不全为零的整系数的方程，可以假定这个方程的首项系数是正的. 我们知道这些多项式的所有可能的整系数是可数的，例如列为 0, +1, −1, +2, −2, +3, −3, …. 全体整系数线性多项式可以排成一个阵列，如

$$\begin{array}{lllllll}
x, & x+1, & x-1, & x+2, & x-2, & x+3, & \cdots \\
-x, & -x+1, & -x-1, & -x+2, & -x-2, & -x+3, & \cdots \\
2x, & 2x+1, & 2x-1, & 2x+2, & 2x-2, & 2x+3, & \cdots \\
-2x, & -2x+1, & -2x-1, & -2x+2, & -2x-2, & -2x+3, & \cdots
\end{array}$$

那么我们可以按照箭头所示的顺序陆续取出上面阵列的对角线元素，把所有线性多项式排成一个单列，其结果是

$$x,\ -x,\ x+1,\ x-1,\ -x+1,\ 2x,\ -2x,\ 2x+1,\ -x-1,\ \cdots.$$

然后我们再求出二次多项式的长方形阵列，这只要把各种二次项

mx^2 添加到上述单列中的每个元素上．由此阵列我们又可得到一个包含所有二次多项式的单列．对高次多项式也如此去做．当对每个次数的多项式都按此法做完之后，结果得到由这些单列组成的阵列，其中第 n 行是包含所有 n 次多项式的单列．再取这个阵列的对角线元素，把它展开，我们就得到包含所有多项式的一个单列．在此单列中，每个多项式都用它的根来代替，并去掉那些重复的根．这个结果就是包含所有整系数多项式的根的单列，这也就是说，所有代数数是可数的．

定理 10 的一个推论是：全体实代数数是可数的．但是康托对角线法证明了(§ 12.3 定理5)所有实数组成的集合是不可数的．因此实数集合大于所有实代数数组成的集合．这个论证对实超越数的存在性给出一个间接证明．我们把这个结果叙述如下：

推论 *不是每个实数都是代数数．*

最初有很多数学家不相信康托的论证，因为这个论证没有给出任何特殊的实超越数．但是现在，他的论证已被广泛接受，并有可能对这个推论给出更明显的证明．

定理 11 *所有代数数组成的集合是一个域．*

我们只须证明，任意两个代数数 $u, v\neq 0$ 的和、积、差、商仍然是代数数．但是所有这些组合都包含在由 u 和 v 生成的复数域的子域 $\mathbf{Q}(u, v)$ 中．因为 u 在 $\mathbf{Q}$ 上是代数的，所以 $\mathbf{Q}(u)$ 是 $\mathbf{Q}$ 的有限扩张；因为 v 在 $\mathbf{Q}(u)$ 上是代数的，所以 $\mathbf{Q}(u, v)$ 是 $\mathbf{Q}(u)$ 的有限扩张．因此根据定理 9，$\mathbf{Q}(u, v)$ 是 $\mathbf{Q}$ 的有限扩张，于是它的每个元素是代数数(定理 8)．　　证毕

如果系数在 F 中的每个多项式方程的根在 F 中，那么称域 F 是代数完全的[①]．在这样的域 F 上每个多项式 $f(x)$ 有一个根 c，因

---

① 有些情况下用"代数闭(algebraically closed)"来代替"代数完全(algebraically complete)"．但考虑到与拓扑学类比，用"代数完全的"这个术语比较好．

此 $f(x)$ 有线性因子 $x-c$. 所以 F 上仅有的一类不可约多项式是线性的，因而代数完全域 F 上每个多项式都可以写成线性因子的乘积(象 § 5.3 公式(11)中所表示那样). 进一步, 除了 F 本身之外, 不可能有 F 的单代数扩张. 于是我们得出结论: 域 F 是代数完全的当且仅当 F 没有真单代数扩张. 代数基本定理(§ 5.3 定理5)断言, 复数域是代数完全的.

定理 12 所有代数数组成的域 A 是代数完全的.

证明 取多项式方程 $x^n+u_{n-1}x^{n-1}+\cdots+u_0=0$, 它的系数 u_i 都是 A 中的代数数. 这些系数生成一个扩张 $K=\mathbf{Q}(u_0, u_1, \cdots, u_{n-1})$, 根据定理 9 的推论 4, 它是有理数域 $\mathbf{Q}$ 的有限扩张. 给定的这个方程的任意复根 r 在域 K 上是代数的, 所以 $K(r)$ 是 K 的有限扩张, 因而也是 $\mathbf{Q}$ 的有限扩张. 根据定理 8, 这个扩张中的元素 r 则在 $\mathbf{Q}$ 上是代数的. 这就意味着根 r 是一个代数数, 所以也就在 A 中, 因此 A 是代数完全的. 证毕

我们现在把域 $\mathbf{Q}$ 嵌入到由所有代数数组成的代数完全域 A 中, 把实数域 $\mathbf{R}$ 嵌入到由所有复数组成的代数完全域 $\mathbf{C}$ 中. 这些结果是下述一般定理的特殊情形, 这个定理指出, 任意域 F 不管怎样都有一个扩张 A, 这个 A 是代数完全的, 并且 A 中每个元素在 F 上是代数的(参看 § 15.1 的附录).

代数数论已发展得很完整. 它主要研究代数数组成的域 K, 它是有理数域 $\mathbf{Q}$ 的有限扩张. 这样的域称为代数数域. 我们下面考虑这个域的算术性质.

习 题

1. 通过求出下列每个代数数所满足的有理系数方程来说明定理 11:

   (a) $\sqrt{2}+\sqrt{-3}$, (b) $\sqrt{-1}+\sqrt[3]{5}$,

   (c) $(\sqrt{7})(\sqrt[3]{2})$, (d) $\dfrac{\sqrt{7}}{1+\sqrt{2}}$,

(e) $u\sqrt{-2}$, 这里 u 满足 $u^3+7u-14=0$.

2. (a) 证明: 如果 u 和 v 分别是($\mathbf{Q}$ 上) m 次和 n 次代数数, 则 $u+v$ 的次数不超过 mn.

(b) $\frac{u}{v}$ 的次数怎样?

(c) 证明: 如果 t 是超越数, u 是代数数, 那么 $t+u$ 和 tu 都是超越数, 对后一种情形, 我们假定 $u\neq 0$.

3. 通过求出下列每个方程的根所满足的有理系数方程来说明定理 12.

(a) $x^2+3x+\sqrt{2}=0$,

(b) $x^2+\sqrt{3}x-\sqrt{-1}=0$,

(c) $x^3-\sqrt{3}x+1+\sqrt[3]{2}=0$,

(d) $x^2+u+2=0$, 这里 u 是方程 $u^3+5u^2-10u+5=0$ 的根.

4. 象定理 10 的证明中所指出的那样, 列出所有二次多项式所组成的单列的前 16 项.

5. 不用定理 10 证明: 固定次数的所有代数数组成的集合是可数的.

6. 证明: 可数域的任意有限扩张是可数的.

7. 证明: 在域 F 的任意可数子域 S 上是代数的所有元素组成的集合 A 是可数的.

8. (a) 证明: 存在一个实数, 它在 $\mathbf{Q}(\pi)$ 上是超越的.

(b) 用习题 7 和 §3.4 的定义证明: 存在可数多个代数无关的实数.

9. 指出在定理 10 的证明中隐含地用到下面的超限算术公式:

(a) 存在 $\mathbf{d}^{n+1}=\mathbf{d}$ 个 n 次多项式.

(b) 存在 $\mathbf{d}+\mathbf{d}+\cdots+\mathbf{d}+\cdots$(到 $\mathbf{d}$ 项) $=\mathbf{d}^2$ 个多项式(所有次数).

*10. (a) 设 u 是任意固定的实数, 通过把 x^j-u^j 因式分解来证明: 存在常数 $N(j)$, 使得不管是否有 $|x-u|<1$, 都有 $|x^j-u^j|\leqslant N|x-u|$.

(b) 设 $f(x)$ 是实系数多项式, u 是任意实数, 证明: 存在一个与 f 和 u 有关的常数 M, 使得不管是否有 $|x-u|<1$, 都有 $|f(x)-f(u)|\leqslant M|x-u|$.

*11. 设实代数数 u 满足 r 次整系数多项式方程 $f(x)=0$. 证明: 如果 m 和 n 是适合 $\left|\frac{m}{n}-u\right|<\frac{1}{Mn^r}$ (其中 M 是习题 10 中的常数), 那么 $f\left(\frac{m}{n}\right)=0$. $\left(\text{提示: 用习题 10, 有}\left|f\left(\frac{m}{n}\right)\right|<\frac{1}{n^r}, \text{而} f\left(\frac{m}{n}\right)\text{是分母为 } n^r \text{ 的有理数.}\right)$

*12. 证明: 如果 u 是实数, 对于 u 可以求出不同有理数的无穷序列

$\frac{m_k}{n_k}$满足$\left|u-\frac{m_k}{n_k}\right|<\frac{1}{kn_k^k}$(对所有的 k)，那么 u 是超越数. $\Big($提示：如果 u 的次数是 r，那么由习题 11 得到，对所有充分大的 k，有 $f\left(\frac{m_k}{n_k}\right)=0.\Big)$

*13. 满足习题 12 的假设条件的数称为刘维尔 (Liouville) 数 (超越数).

(a) 证明：$\sum_{k=1}^{\infty}10^{-k!}=0.110001\cdots$是刘维尔数.

(b) 再列出两个别的刘维尔数.

§14.7 高斯整数

高斯整数是其分量 a，b 都为整数的复数 $a+bi$. 任意这样的高斯整数满足整系数的首一多项式方程 $\alpha^2-2a\alpha+(a^2+b^2)=0$，因此它是代数数. 两个高斯整数的和、差、积仍然是高斯整数，因此全体高斯整数构成整环 $\mathbf{Z}[i]$. 在这个整环中可以考虑可除性和分解成素(不可约)因子的问题.

对任意复数 σ，引进"范数"(是整数，或者不是)是方便的. 设 $\sigma=r+si$，则范数 $N(\sigma)$ 是 σ 与它的共轭复数 $\sigma^*=r-si$ 的乘积：

$$N(\sigma)=\sigma\sigma^*=(r+si)(r-si)=r^2+s^2. \tag{10}$$

这个范数永远是非负的，并且是 σ 绝对值的平方. 对任意两个复数 σ 和 τ，我们有

$$N(\sigma\tau)=N(\sigma)N(\tau). \tag{11}$$

这个等式意味着，对应$\sigma\longmapsto N(\sigma)$保持乘积，换句话说，它是由非零的数 σ 组成的乘法群到实数的乘法群上的同态映射. 特别，高斯整数的范数是(有理)整数①.

现在回忆一下包含可除性的一般概念 (§ 3.6). $\mathbf{Z}[i]$的单位

① 有理整数即普通整数，定义见下节. ——译者注.

是这样的高斯整数：$\alpha\neq0$，它的倒数 α^{-1} 也是一个高斯整数. 那么有 $\alpha\alpha^{-1}=1$，所以 $N(\alpha\alpha^{-1})=N(\alpha)N(\alpha^{-1})=1$，因而单位 α 的范数一定是 $N(\alpha)=1$. 由(10)式可以看出，$\mathbf{Z}[i]$的单位只能是 ±1 和 $\pm i$. 如果 $\mathbf{Z}[i]$中两个高斯整数可彼此整除，则称这两个整数在 $\mathbf{Z}[i]$中相伴. 因此在 $\mathbf{Z}[i]$中 α 的相伴只能是 $\pm\alpha$ 和 $\pm i\alpha$.

有理素数 5 在 $\mathbf{Z}[i]$中有四种不同的分解

$$5=(1+2i)(1-2i)=(2i-1)(-2i-1)$$
$$=(2+i)(2-i)=(i-2)(-i-2). \tag{12}$$

这些分解本质上是一样的，例如，$2+i=i(1-2i)$和 $2-i=-i(1+2i)$，其他各种情形中相应的因子也都是相伴. (12)中的每个因子都是素的①. 例如，如果 $2+i$ 可以分解为 $2+i=\alpha\beta$，那么 $N(2+i)=5=N(\alpha)N(\beta)$，所以$N(\alpha)$(或者$N(\beta)$)等于 1，因此 α (或者 β)是单位. (12)式的各个因子本质上只给出 5 的一个因子分解，因为在任意分解 $5=\gamma\delta$ 中，$N(5)=25=N(\gamma)N(\delta)$，所以每个不是单位的因子一定有范数 5. 通过试验我们发现，范数为 5 的所有高斯整数就是(12)式中写出的那些.

另一方面，有理素数 3 在 $\mathbf{Z}[i]$ 中是素的. 假定 $3=\alpha\beta$，则 $N(\alpha)N(\beta)=9$，于是 $N(\alpha)|9$. 如果 $N(\alpha)=1$，则 α 是单位；如果 $N(\alpha)=N(a+bi)=3$，则 $a^2+b^2=3$，不可能有整数 a 和 b 使这个等式成立. 因此在高斯整数环中 3 没有真因子 α.

高斯整数的唯一因子分解定理可以通过构造除法算式来证明，这个算式类似于对普通整数和多项式用过的除法算式.

定理 13 对于给定的高斯整数 α 和 $\beta\neq0$，存在高斯整数 γ 和 ρ 适合

$$\alpha=\beta\gamma+\rho,\quad N(\rho)<N(\beta). \tag{13}$$

① 即不可分解的，其定义见后，它与有理素数即普通素数的定义不同. ——译者注.

证明 我们从商$\frac{\alpha}{\beta}=r+si$出发，选取$r'$和$s'$是与有理数$r$和$s$最接近的整数．那么

$$\frac{\alpha}{\beta}=(r'+s'i)+[(r-r')+(s-s')i]=\gamma+\sigma,\quad \gamma=r'+s'i,$$

这里$|r-r'|\leqslant\frac{1}{2}$, $|s-s'|\leqslant\frac{1}{2}$,所以

$$N(\sigma)=(r-r')^2+(s-s')^2\leqslant\frac{1}{4}+\frac{1}{4}<1.$$

上述方程现在可写成$\alpha=\beta\gamma+\beta\sigma$,其中$\alpha$, $\beta\gamma$都是高斯整数,因此$\beta\sigma$也是高斯整数,而且这里$N(\beta\sigma)=N(\beta)N(\sigma)<N(\beta)$.

证毕

引理 1 两个高斯整数α_1和α_2有最大公因子δ,它是一个可表示成$\delta=\beta_1\alpha_1+\beta_2\alpha_2$形式的高斯整数，其中$\beta_1$和$\beta_2$都是高斯整数．

证明 通过辗转相除,我们可以构造欧几里得算法,它很象有理整数情形(§ 1.7)．(13)式的逐次余数ρ的范数越来越小，因此这种算法最终会结束．最后的非零余数就是我们所要求的最大公因子．

证毕

另一个证明是从环$\mathbf{Z}[i]$中的α_1和α_2生成的理想(α_1, α_2)出发．在这个理想的全体元素中间选择一个范数最小的元素δ，并象(13)式那样,写$\alpha_1=\delta\gamma_1+\rho_1$, $\alpha_2=\delta\gamma_2+\rho_2$．这些系数$\rho_i$都在理想中,并且它们的范数小于$\delta$的范数,因此一定是零．所以$\alpha_1=\delta\gamma_1$, $\alpha_2=\delta\gamma_2$，于是δ是公因子．因为δ在这个理想中，所以它有形式$\delta=\beta_1\alpha_1+\beta_2\alpha_2$，因此它是$\alpha_1$和$\alpha_2$的每个公因子的倍数．所以$\delta$就是我们所要求的最大公因子．

高斯整数分解的其余部分可同有理整数的情形（§ 1.7 和 § 1.8)及多项式的情形(§ 3.5 和 § 3.8)完全一样地进行处理，因

此我们这里只叙述这重要的步骤. 一个高斯整数π, 如果它不是0也不是单位, 而且它在$\mathbf{Z}[i]$中的因子只能是单位和π的相伴, 则称π是素的. 我们可以证明

引理 2 如果π是素的, 那么由$\pi|\alpha\beta$可推出$\pi|\alpha$或者$\pi|\beta$.

定理 14 每个高斯整数α可以表示成素高斯整数的乘积$\alpha=\pi_1\cdots\pi_n$. 这个表达式实质上是唯一的, 所谓实质上是唯一的是指, 任意其他α表示成素高斯整数之积的分解式有相同的因子个数, 并可重新排列使得相应位置的因子是相伴.

为了适当地推广这些概念, 我们首先研究一下高斯整数所满足的不可约多项式方程. 如果$\alpha=a+bi$是高斯整数, 而不是有理整数, 那么$b\neq0$, 并且α一定满足一个二次不可约方程. 这就是

$$[x-(a+bi)][x-(a-bi)]=x^2-2ax+(a^2+b^2)=0.$$

它是以有理整数为系数的首一不可约方程. 反过来, 可以证明[①], 如果数$r+si$在域$\mathbf{Q}(i)$中满足一个整系数首一不可约方程, 那么这个数是高斯整数. 这就给出

定理 15 在域$\mathbf{Q}(i)$中的一个数是高斯整数当且仅当它在$\mathbf{Q}$上所满足的首一不可约方程是以整数为系数.

习　题

1. 把下列各高斯整数分解成素因子乘积:

$$5,\ 3+i,\ 6i,\ 11,\ 1-7i.$$

2. 求出下列每对高斯整数α_1和α_2的最大公因子, 并把它表示成$\beta_1\alpha_1+\beta_2\alpha_2$的形式:

(a) $3+6i$ 和 $12-3i$.

(b) $5+3i$ 和 $13+18i$.

3. 求出13的所有可能的因子分解(分解成素高斯整数之积), 并证明: 任意两个分解仅差相伴.

---

① 在下一节中(定理16), 对稍微更一般的情形给出证明.

4. 证明：由高斯整数构成的每个理想都是主理想.

5. (a) 用欧几里得算法证明引理 1.

(b) 证明引理 2.

6. 由引理 2 证明定理 14.

7. (a) 证明：有理素数 p 在 $\mathbf{Z}[i]$ 中是素的当且仅当方程 $x^2+y^2=p$ 没有整数解 x 和 y.

(b) 证明：任意形为 $p=4n+3$ 的有理素数在 $\mathbf{Z}[i]$ 中是素的.

*8. (a) 证明：商环 $\mathbf{Z}[x]/(p, x^2+1)$ 既与 $\mathbf{Z}[i]/(p)$ 同构，也与 $\mathbf{Z}_p[x]/(x^2+1)$ 同构.

(b) 证明：$\mathbf{Z}[i]/(p)$ 是整环当且仅当 p 在 $\mathbf{Z}[i]$ 中是素的；$\mathbf{Z}_p[x]/(x^2+1)$ 是整环当且仅当 $x^2\equiv-1(\bmod p)$ 在 $\mathbf{Z}$ 中没有解.

(c) 假设模 p 乘法群是循环群(§15.3 定理6)，证明：如果 $p=4n+1$，那么 $x^2\equiv-1(\bmod p)$ 在 $\mathbf{Z}$ 中有解.

(d) 证明：$p=4n+1$ 在 $\mathbf{Z}[i]$ 中不可能是素的.

习题 9～13 都指的是由数 $a+b\sqrt{-2}$（其中 a 和 b 是整数）构成的整环 $\mathbf{Z}[\sqrt{-2}]$.

9. 定义范数 $N(a+b\sqrt{-2})=a^2+2b^2$，并列出它的性质.

10. 证明整环 $\mathbf{Z}[\sqrt{-2}]$ 中的除法算式.

11. 证明整环 $\mathbf{Z}[\sqrt{-2}]$ 中最大公因子的存在性.

12. 叙述并证明 $\mathbf{Z}[\sqrt{-2}]$ 中的唯一因子分解定理.

13. 在 $\mathbf{Z}[\sqrt{-2}]$ 中，把下列各数因子分解：$5, 1+3\sqrt{-2}, 2+\sqrt{-2}$.

14. (a) 在 $\mathbf{Z}[\sqrt{2}]$ 中求一个不同于 ±1 的单位.

(b) 证明：在 $\mathbf{Z}[\sqrt{2}]$ 中存在无穷多个不同的单位.（提示：用一个单位的幂.）

§14.8 代数整数

一般地，如果代数数 u 满足的有理数域上首一不可约方程是以整数为系数，即

$$p(u)=a_0+a_1u+\cdots+a_{n-1}u^{n-1}+u^n=0,\ a_i \text{ 是整数} \qquad (14)$$

其中 $p(x)$ 在 $\mathbf{Q}$ 上是不可约的，那么称 u 是代数整数. 有理数 $\frac{m}{n}$ 所

满足的不可约方程刚好是线性方程 $x-\frac{m}{n}=0$. 所以一个有理数是代数整数当且仅当它是一个普通意义下的整数. 于是 $\mathbf{Z}$ 的（普通）整数称为有理整数，以便同其他代数整数相区别. 如果代数数 $u\neq0$, u 和 u^{-1} 都是代数整数，那么称 u 是单位.

在检验给定的代数数是否是代数整数时，不一定要借助于不可约方程，而依赖下面的定理.

定理 16 一个数是代数整数当且仅当它在 $\mathbf{Q}$ 上满足一个整系数首一多项式方程.

证明 假设 u 是某个整系数首一多项式 $f(x)$ 的根，在 $\mathbf{Q}$ 上，u 还满足一个不可约多项式 $p(x)$，它可以取为整系数的. 这些系数的任意公因子可被消去，因此我们可以假定 $p(x)$ 的系数的最大公因子是 1. 这也就是说，在 § 3.9 的意义下，$p(x)$ 是由所有整系数多项式组成的整环 $\mathbf{Z}[x]$ 中的本原多项式. 因为已知的多项式 $f(x)$ 是首一多项式，所以也是本原的. 由定理 2 我们知道，以 u 为根的多项式 $f(x)$ 在 $\mathbf{Q}[x]$ 中一定被 u 所满足的不可约多项式 $p(x)$ 整除，于是 $f(x)=q(x)p(x)$. 因为 f 和 p 都是本原的，所以根据 § 3.9 引理 3 可以断言，商式 $q(x)$ 也具有整系数. 那么 $f(x)$ 的首项系数 1 是 q 和 p 的首项系数之积，因此 $\pm p(x)$ 是首一多项式，根据定义 (14)，这意味着 u 是代数整数.

一个数虽然看起来不象代数整数，但实际上可能是代数整数，例如，$u=\frac{1+\sqrt{5}}{2}$ 看起来象一个分数，但它满足方程

$$\left(x-\frac{1+\sqrt{5}}{2}\right)\left(x-\frac{1-\sqrt{5}}{2}\right)=x^2-x-1=0.$$

这个方程是首一整系数方程. 这就暗示我们系统地找出在二次域中是代数整数的那些数. 有理数域 $\mathbf{Q}$ 上的任意二次域 K 可以表示成单代数扩张 $K=\mathbf{Q}(\sqrt{d})$. 不失一般性，我们可以假定 d 是一

个整数且无平方因子(1除外)．这就是我们所要考虑的情形:

定理 17 如果 $d\neq 1$ 是无平方因子的整数，那么在 $d\equiv 2$ 或 $d\equiv 3 \pmod 4$ 的情形下，$\mathbf{Q}(\sqrt{d})$ 中的代数整数是形为 $a+b\sqrt{d}$（其中系数 a 和 b 是有理整数）的数．但是当 $d\equiv 1 \pmod 4$ 时，$\mathbf{Q}(\sqrt{d})$ 中的代数整数是形为 $a+b\dfrac{1+\sqrt{d}}{2}$（其中 a 和 b 为有理整数）的数．

证明 作为预备知识，我们注意到，$a\equiv 1 \pmod 2$ 意味着 $a=1+2r$，因此 $a^2=1+4r+4r^2\equiv 1 \pmod 4$．换句话说，

$$a\equiv 1 \pmod 2 \quad 可推出\ a^2\equiv 1 \pmod 4, \tag{15}$$

$$a\equiv 0 \pmod 2 \quad 可推出\ a^2\equiv 0 \pmod 4, \tag{16}$$

所以一个平方数总同余于 0 或 1, mod 4.

$\mathbf{Q}(\sqrt{d})$ 中任意数 u 可表示成 $u=\dfrac{a+b\sqrt{d}}{c}$，这里整数 a, b, c 无公因子．为了排除有理数的平凡情形，我们假定 $b\neq 0$．那么 u 所满足的二次首一不可约方程是

$$\left(x-\frac{a+b\sqrt{d}}{c}\right)\left(x-\frac{a-b\sqrt{d}}{c}\right)$$

$$=x^2-\frac{2a}{c}x+\frac{a^2-db^2}{c^2}=0. \tag{17}$$

如果 u 是代数整数，那么这些系数 $\dfrac{2a}{c}$ 和 $\dfrac{a^2-db^2}{c^2}$ 也一定是整数．所以 $\dfrac{4a^2}{c^2}$, $\dfrac{4a^2-4db^2}{c^2}$ 和 $\dfrac{4db^2}{c^2}$ 都必是整数，所以 $c\mid 2a$, $c^2\mid 4db^2$．因为已假定 d 不含平方因子，所以包含在 c 中的任意素数 $p\neq 2$ 一定同时整除 a 和 b^2，这与 a, b, c 无公因子（± 1 除外）的假定相矛盾．由于类似的理由，$4\mid c$ 是不可能的，所以只能选取 $c=1$ 和 $c=2$.

现在考虑 $d\equiv 2$ 或 $d\equiv 3 \pmod 4$ 的情形，取 $c=2$．在这种情形下，(17)式最后的系数 $\dfrac{a^2-db^2}{4}$ 一定是整数，于是 $a^2\equiv db^2 \pmod 4$.

如果$b\equiv 1 \pmod 2$，则$b^2\equiv 1 \pmod 4$，并且$a^2\equiv db^2\equiv 2$或$3 \pmod 4$，这与法则(15)和(16)相矛盾。如果$b\equiv 0 \pmod 2$，则$a^2\equiv 0 \pmod 4$，因而$a\equiv 0 \pmod 2$，所以a, b, c有公因子2. 无论哪一种情形，我们都得出$c=1$，所以$\mathbf{Q}(\sqrt{d})$的所有代数整数是形为$a+b\sqrt{d}$的数. 反过来，这种形式的数所满足的首一方程(17)具有整系数.

剩下的情形$d\equiv 1 \pmod 4$可以类似处理，除了可能出现$a\equiv b\equiv 1 \pmod 2$的情形以外，其余都类似.

推论 Q上任意二次域中，所有代数整数组成的集合是一个整环.

证明 定理17中所表示的代数整数的和、差、积仍然是这种形式的代数整数. 证毕

下面的任务就是把这个推论推广到任意代数数域上.

习 题

1. 证明：每个单位根是代数整数.

2. (a) 求$\mathbf{Q}(\omega)$中的所有代数整数和所有单位，这里ω是复三次单位根.

   (b) 证明：$\mathbf{Q}(\omega)$中每个单位都是单位根.

3. 完成定理17的第二种情形($d\equiv 1 \pmod 4$)的证明.

4. (a) 证明：任意代数数可以写成商$\frac{u}{b}$，这里u是代数整数，b是有理整数(即$\mathbf{Z}$中的整数).

   (b) 证明：代数数组成的任意域K是由K中的所有代数整数组成的整环的商域.

*5. 求出$\mathbf{Q}(\sqrt{2}, i)$中的所有代数整数.

§14.9 代数整数的和与积

这一节来证明下述结果：

定理18 所有代数整数的集合是一个整环.

下面是定理的直接推论.

推论 *在由代数数组成的任意域 K 中，全体代数整数构成一个整环.*

定理 18 的一个启发性的证明依赖于对代数整数生成的加法群的分析. 如果 $v_1, \cdots, v_n$ 是任意代数数，我们令 $G=[v_1, \cdots, v_n]$ 表示在由全体复数组成的加法群中由这些代数数生成的子群①. 这个群只是由可表示成形式

$$u=a_1v_1+a_2v_2+\cdots+a_nv_n \quad (a_i \text{ 是有理整数}) \tag{18}$$

的所有数组成. 回忆一下，在由 v 生成的加法循环子群中，自然倍数 $av=a\times v$ 就是 v 的"幂".

引理 1 *群 $G=[v_1, \cdots, v_n]$ 的任意子群 S 也可以由 n 个或更少一些数生成.*

证明 对每个下标 k，设 G_k 是由 G 的后 $n-k+1$ 个生成元生成的子群 $[v_k, \cdots, v_n]$，所以 G_k 是由所有形为 $a_kv_k+\cdots+a_nv_n$ 的和构成. 在 G_k 的位于给定子群 S 的元素中间选取一个元素

$$w_k=c_kv_k+c_{k+1}v_{k+1}+\cdots+c_nv_n, \tag{19}$$

使其中第一个系数 c_k 有最小的正值 (这是可能的). (如果对每个元素，v_k 的系数都是零，则置 $w_k=0$.) 如果 $w=b_kv_k+\cdots$ 是 G_k 中 S 的任意其他元素，它的第一个系数 b_k 可以写成 $b_k=q_kc_k+r_k$，具有一个非负余数 $r_k<c_k$. 那么这个差 $w-q_kw_k=r_kv_k+\cdots$ 就在群 G_k 和 S 中，并有非负的第一个系数 r_k，它小于最小的正值 c_k，因此 $r_k=0$，于是 G_k 中任意 S 的元素 w 给出 G_{k+1} 中的一个元素 $w'=w-q_kw_k$.

这样选取的 n 个元素 $w_1, \cdots, w_n$ 生成整个群 S，这是因为在 S 中任意给定元素 w，我们就可以找到 q_1，使 $w-q_1w_1$ 只依赖于

① 这个加法群有时称为 $\mathbf{Z}$-模，因为它的元素可以用 $\mathbf{Z}$ 中的标量来乘.

$v_2, \cdots, v_n$，然后又可找到某个 q_2，使 $w-q_1w_1-q_2w_2$ 只依赖于 $v_3, \cdots, v_n$，等等. 最后有 $w=\sum q_iw_i$. 证毕

引理 2 数 u 是代数整数当且仅当由 u 的所有幂 $1, u, u^2, u^3, \cdots$ 生成的加法群可以由有限个元素生成.

证明 如果 u 是代数整数，则它满足整系数的 n 次首一方程(14). 这个方程把 u^n 表示成群 $G=[1, u, \cdots, u^{n-1}]$ 中的一个元素，这个群是由 u 的低于 n 次的幂生成的. 通过迭代，上述同一个方程可以用来把 u 的任意高次幂表示成这个群中的一个元素. 所以 u 满足引理 2 的准则.

反过来，假定由 $1, u, u^2, \cdots$ 生成的群 G 可以由 G 的任意 n 个数 $v_1, \cdots, v_n$ 生成. u 与 G 的任意元素 $\sum a_ju^j$ 的乘积仍然是 G 的元素 $\sum a_ju^{j+1}$，所以每个乘积 uv_i 一定在 G 中，于是它一定可以按照生成元表示成 $uv_i=\sum\limits_j a_{ij}v_j$，这里 a_{ij} 是整数. 这些表达式给出如下形式的 n 个 v 的齐次方程

$$
\begin{gathered}
(a_{11}-u)v_1+a_{12}v_2+\cdots+a_{1n}v_n=0, \\
a_{21}v_1+(a_{22}-u)v_2+\cdots+a_{2n}v_n=0, \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
a_{n1}v_1+a_{n2}v_2+\cdots+(a_{nn}-u)v_n=0.
\end{gathered}
$$

这组方程有一组不全为零的解 $v_1, v_2, \cdots, v_n$，所以系数矩阵的行矢量一定是线性相关的(§ 7.7 定理 13 的推论). 这个系数矩阵可以写成 $A-uI$，这里 $A=(a_{ij})$. 因为它是奇异的，所以它的行列式是零，于是

$$
|A-uI|=(-1)^nu^n+b_{n-1}u^{n-1}+\cdots+b_n=0, \tag{20}
$$

这里系数 b_i 是整数 a_{ij} 的某一多项式，因而它们都是整数. 这个方程(20)意味着① u 是代数整数，正如引理中所要求的.

---

① 注意，根据第十章的意义(20)式就是 A 的特征多项式.

引理 2 的结论可以重述如下：

推论 如果代数数 u 的所有正次幂都在由一组有限个数 $y_1, \cdots, y_n$ 生成的加法群中，那么 u 是代数整数.

证明 由 $1, u, u^2, \cdots$ 生成的群 S 是由 $1, y_1, \cdots, y_n$ 生成的群的一个子群. 因此根据引理 1，这个子群 S 可以由它的有限多个元素生成，因此根据引理 2，数 u 是代数整数. 证毕

现在回来证明定理 18. 如果 u 和 v 是代数整数，我们应指出 $u+v$ 和 uv 都是代数整数. 这个假设条件意味着所有的幂 u^k 和 v^k 分别可以按照有限多个幂 $1, u, \cdots, u^{n-1}$ 和 $1, v, \cdots, v^{r-1}$ 来表示. 所以每个幂 $(uv)^k = u^k v^k$ 和 $(u+v)^k$ 在由乘积 $1, u, uv, uv^2, \cdots, u^{n-1}v^{r-1}$ 生成的加法群中. 根据推论得出，uv 和 $u+v$ 是代数整数，这正是定理所要求的.

习 题

1. 通过列出适当的整系数首一方程，明显地证明下列各数都是代数整数：

(a) $\sqrt{2}+\sqrt{3}$, (b) $i+\omega$,

(c) $\sqrt{7}+\dfrac{1+\sqrt{5}}{2}$.

2. (a) 证明：如果数 $v_1, \cdots, v_n$ 在 $\mathbf{Q}$ 上线性无关，那么在 $G=[v_1, \cdots, v_n]$ 中的任意有有限指数的子群 S 也可以由 n 个线性无关的数 $w_1, \cdots, w_n$ 生成.

(b) 证明：任意这样的子群 S 与整个群 G(群)同构.

3. 如果数 $v_1, \cdots, v_n$ 在 $\mathbf{Q}$ 上线性无关，指出怎样用引理 1 中对于 $G=[v_1, \cdots, v_n]$ 的子群 S 求出的那组基底来计算 S 在 G 中的指数.（提示：求出 S 的每个陪集的代表元.）

*4. 证明：群 $G=[v_1, \cdots, v_n]$ 没有不同子群的无限升链. 也就是证明：给出子群的无穷序列 $S_1 \leqslant S_2 \leqslant S_3 \leqslant \cdots \leqslant G$，则存在一个下标 m，使得 $S_m = S_{m+1} = S_{m+2} = \cdots$.（提示：把引理 1 应用到全体群 S_k 的并.）

5. (a) 证明：包含在普通整数环 **Z** 中的每个模是 **Z** 的一个理想.

(b) 列出一个包含在高斯整数的整环 $\mathbf{Z}[i]$中不是 $\mathbf{Z}[i]$中的理想的模.

*6. 证明：如果代数数 u 满足一个首一多项式方程，这个方程的其他系数都是代数整数，那么 u 也是代数整数.

§14.10 二次代数整数的因子分解

为了说明代数整数因子分解理论，我们更详细地考虑最简单的情形，即二次代数整数的因子分解. 也就是说，我们考虑 $\mathbf{Q}(\sqrt{d})$的代数整数(象定理 17 中所描述的那种代数整数)的因子分解. 为此目的所用的基本工具是范数概念.

范数的公式依赖于域，但是范数的概念在所有情况下，甚至对于高次代数数域，都是一样的. 本质上，范数是通过域的自同构来定义的. 根据定理 6，二次域 $\mathbf{Q}(\sqrt{d})$有一个自同构 $u=a+b\sqrt{d}\longleftrightarrow\bar{u}=a-b\sqrt{d}$，它把每个数映射到它的共轭数 $\bar{u}$.

定义 $\mathbf{Q}(\sqrt{d})$的数 $u=a+b\sqrt{d}$ 的范数 $N(u)$ 是 u 和它的共轭 $\bar{u}$ 的乘积 $u\bar{u}$:

$$N(u)=u\bar{u}=(a+b\sqrt{d})(a-b\sqrt{d}). \tag{21}$$

因为对应 $u\longleftrightarrow\bar{u}$ 是同构，$\overline{uv}=\bar{u}\cdot\bar{v}$，所以

$$N(uv)=N(u)N(v). \tag{22}$$

于是把这个域中的代数整数的任意分解 $w=uv$ 转换成有理整数 $N(w)$的分解 $N(w)=N(u)N(v)$.（代数整数的范数是有理整数，见习题 1.）

范数的性质基本上依赖于 d 是正的还是负的，也就是依赖于 $\mathbf{Q}(\sqrt{d})$ 是实二次域还是复二次域. 如果 $d<0$，则 $N(u)$ 就是 $|u|^2$，即 u 的绝对值平方，除了 $u=0$ 外，它是正的. 如果 $d>0$，则 $N(u)=a^2-b^2d$ 可以是正的也可以是负的. 这个差别出现在 $\mathbf{Q}(\sqrt{d})$的单位的群 U 中，正如我们将看到的

引理1 代数整数 $u\in\mathbf{Q}(\sqrt{d})$ 是单位当且仅当 $N(u)=\pm1$.

证明 显然，$N(1)=1$；此外，$N(u)$一定是有理整数. 因此，如果对某一个其他代数整数 $v\in\mathbf{Q}(\sqrt{d})$ 有 $uv=1$，那么有 $N(u)\cdot N(v)=N(uv)=1$，因此 $N(u)=\pm1$. 反过来，如果 $N(u)=u\bar{u}=\pm1$，那么 $u(\pm\bar{u})=1$，于是 u 是 $\mathbf{Q}(\sqrt{d})$ 的单位.

类似的论证可应用到一般代数数域上.

把引理1和定理17结合起来，我们可以确定任意复二次域 $\mathbf{Q}(\sqrt{-d})$ ($d>0$，无平方因子的整数)的全部单位. 那么 $\mathbf{Q}(\sqrt{-d})$ 的代数整数具有形式 $u=m+n\alpha$ $(m, n\in\mathbf{Z})$，这里

$$\alpha=\begin{cases}\sqrt{-d}, & \text{当 } d\not\equiv3 \pmod 4\\ \dfrac{1+\sqrt{-d}}{2}, & \text{当 } d\equiv3 \pmod 4\end{cases}$$

相应地，u 的范数满足

$$N(u)=\begin{cases}m^2+n^2d, & \text{当 } d\not\equiv3 \pmod 4\\ \left(m+\dfrac{n}{2}\right)^2+\dfrac{n^2d}{4}, & \text{当 } d\equiv3 \pmod 4\end{cases}$$

当 $d\not\equiv3 \pmod 4$ 且 $d>1$ 时，$m^2+n^2d\leqslant1$ 只有当 $m=\pm1$，$n=0$ 时才有可能. 同样地，如果 $d\equiv3 \pmod 4$ 且 $d>3$，那么 $d\geqslant7$，并且 $N(u)\geqslant\dfrac{7n^2}{4}>1$，除非 $n=0$. 因此又一次说明 $\mathbf{Q}(\sqrt{-d})$ 的单位只能是 ±1. 这就证明了

定理19 *存在不同于 ±1 的单位的复二次域只能是 $\mathbf{Q}(\sqrt{-1})$ 和 $\mathbf{Q}(\sqrt{-3})$.*

$\mathbf{Q}(\sqrt{-1})$ 的单位是 ±1 和 $\pm i$；$\mathbf{Q}(\sqrt{-3})$ 的单位是 $\omega=\dfrac{1+\sqrt{-3}}{2}$ 的各次幂，ω 是六次本原单位根.

实二次域有无穷多个单位. 例如，$1+\sqrt{2}$ 是 $\mathbf{Q}(\sqrt{2})$ 的单位，

因为$N(1+\sqrt{2})=-1$. 因此$1+\sqrt{2}$的所有次幂$(1+\sqrt{2})^{\pm k}$都是单位.

虽然对于很多二次代数整数的环，分解成素因子的因子分解是唯一的，但在$\mathbf{Q}(\sqrt{-5})$中情况并非如此. 例如，考虑数 6 的因子分解：

$$6=2\cdot3=(1+\sqrt{-5})(1-\sqrt{-5}). \tag{23}$$

如果$\mathbf{Q}(\sqrt{-5})$的两个代数整数u和v满足$uv=6$，那么$N(u)N(v)=N(6)=36$. 于是 6 的真因子u的范数将是$2^2 3^2$的真因子，所以只有$N(u)=2, 3, 4, 6, 9, 12, 18$几种情况需要研究. 因为在这些情况中，$N(v)$分别为 18, 12, 9, 6, 4, 3, 2，所以只须考虑$u=2, 3, 4, 6$等情形. 由$N(m+n\sqrt{-5})=m^2+5n^2$容易看出，所有可能的因子都已列在(23)式中.

在上述例子中，我们可以考虑用理想的乘积(如§13.4中所述)代替数的乘积，以此来补救唯一因子分解定理. 我们发现主理想(2), (3), $(1+\sqrt{-5})$和$(1-\sqrt{-5})$都不是素理想. 相关的素理想是$P=(2, 1+\sqrt{-5})$, $Q=(3, 1+\sqrt{-5})$，这是用它们在$\mathbf{Z}(\sqrt{-5})$中的基底来描述的. 这些理想不是主理想，把它们平方起来：

$$P^2=(4, 2+2\sqrt{-5}, 6)=(2)$$
$$Q^2=(9, 3+3\sqrt{-5}, 6)=(3)$$

这表明(2)和(3)不是素理想.

为证明P是$\mathbf{Z}[\sqrt{-5}]$中的素理想，我们注意，$(m+n\sqrt{-5})\in P$当且仅当$m+n\equiv0 \pmod 2$. 因此$\mathbf{Z}[\sqrt{-5}]/P$只包含两个元素，它是域$\mathbf{Z}_2$. 因此根据§13.3 定理 6，P是素理想. 类似地，$\mathbf{Z}[\sqrt{-5}]/Q$是$\mathbf{Z}_3$，所以Q也是素理想.

总之，我们证明了$\mathbf{Z}[\sqrt{-5}]$的理想(6)具有分解成素理想的唯一因子分解$(6)=P^2Q^2$.

我们在整环$\mathbf{Z}[\sqrt{-5}]$中推导出的这种理想的唯一分解仅仅

是用以说明，理想的概念怎样可以系统地用来建立代数整数环上的唯一因子分解定理，而通常的因子分解在这个整环上是不唯一的．通过进一步推理，我们可以建立“理想论基本定理”：在代数数域K中由所有代数整数组成的整环D中，每个理想除次序外可以唯一地表示成素理想的乘积．特别是，整环的每个代数整数u确定一个主理想(u)，在上述意义下，它有唯一的因子分解．

习　题

1．(a) 证明：在任意二次域中，代数整数的范数是有理整数．

(b) 证明：如果$u=a+b\sqrt{d}$不是有理数，那么$N(u)$是u所满足的首一不可约多项式方程的常数项．

2．求$\mathbf{Q}[\sqrt{-7}]$中的全部单位．

3．证明：二次域$\mathbf{Q}(\sqrt{-d})$（其中$d>0$）中单位的个数是有限的，并证明：每个单位是单位根．

*4．证明：任意给定的代数数域中的全体单位根构成循环群．

5．叙述并证明$\mathbf{Z}[\omega]$的除法算式，这里$\omega=\dfrac{-1+\sqrt{-3}}{2}$．（提示：任意$\beta$的整数倍数把复平面分割成无穷多个等边三角形．）

*6．设D是任意整环，在这个整环中范数$N(\alpha)$定义如下：(i)当$\alpha\neq 0$，$N(\alpha)$是正整数；(ii)$N(\alpha\beta)=N(\alpha)N(\beta)$；(iii)给定$\alpha$和$\beta\neq 0$，存在$\gamma$和$\zeta$使得$\alpha=\beta\gamma+\zeta$，$N(\zeta)<N(\beta)$．

(a) 证明：D是唯一因子分解整环．

(b) 证明：D中每个理想是主理想．

第十五章　伽罗瓦理论

§15.1　方程的根域

历史上很多代数学家试图用明显的公式求解实多项式方程和(以后的)复多项式方程．在他们的努力下，求出了一般的二次、三次和四次方程的“根式解”，这些求解公式我们已在第五章推导过．但是对于五次方程，多次想得到类似的求解公式，结果都失败了．

这是为什么？其原因最后被伽罗瓦(Évariste Galois)发现，他指出一个方程有根式解当且仅当与它相联系的自同构群在纯群论意义下是“可解”的．这里所说的自同构是指由这个方程的所有根生成的扩域的，使方程的所有系数保持固定的那些自同构．在最后这一章里，我们从讨论给定域 F 上的已知多项式 $p(x)$ 的所有根生成的扩域开始，按照现代的形式介绍伽罗瓦最本质的论证．这个域就是所谓 $p(x)$ 的“根域”，① 现在我们正式地给出它的定义．

定义　F 的扩张 N 如果满足下列条件则称为系数在 F 中的 n ($\geqslant 1$) 次多项式 $f(x)$ 的根域：(i) $f(x)$ 在 N 中可以分解成线性因子 $f(x)=c(x-u_1)\cdots(x-u_n)$；(ii) N 是 F 上添加 $f(x)$ 的全部根而生成的，即 $N=F(u_1,\cdots,u_n)$．

如果 $f(x)=ax^2+bx+c(a\neq0)$ 是 F 上的二次多项式，它有共轭根② $u_j=\dfrac{-b\pm\sqrt{b^2-4ac}}{2a}$，$j=1, 2$，由 $f(x)=0$ 的一个根 u_1 生成的 F 的单扩张 $K=F(u_1)\cong F[x]/(f(x))$ 已经是 f 在 F 上的根

---

① 有的代数书中称这个域为分裂域．——译者注．

② 当然，多项式 $f(x)$ 的根是指满足 $f(x)=0$ 的数 x，这个 x 也称为 $f(x)$ 的零点．

域. 这是因为$u_2=\frac{c}{au_1}$, 因此$f(x)$在域$K=F(u_1)$中可以分解成线性因子$f(x)=a(x-u_1)(x-u_2)$.

可是, 对于三次不可约多项式, 这个结论一般来说是不正确的. 例如, $\mathbf{Q}$上不可约多项式x^3-5的根域N是$\mathbf{Q}(\sqrt[3]{5},\omega\sqrt[3]{5},\omega^2\sqrt[3]{5})=\mathbf{Q}(\sqrt[3]{5},\omega)$, 其中$\omega=\frac{-1+i\sqrt{3}}{2}$是复三次单位根. 由5的实三次根生成的有理数域的实扩张$\mathbf{Q}(\sqrt[3]{5})\cong\mathbf{Q}[x]/(x^3-5)$在$\mathbf{Q}$上的次数是3, 而包含5的所有三次根的$\mathbf{Q}$的最小扩张是$N=\mathbf{Q}(\sqrt[3]{5},\omega)$. 因为$\omega$满足分圆方程$\omega^2+\omega+1=0$, 所以域$N$在$\mathbf{Q}(\sqrt[3]{5})$上的次数是2. 当我们把$x^3-5$的根域$N$看作$\mathbf{Q}$上的一个矢量空间时, 它就有基底$(1,\ \sqrt[3]{5},\ \sqrt[3]{25},\ \omega,\ \omega\sqrt[3]{5},\ \omega\sqrt[3]{25})$, 于是它是$\mathbf{Q}$的六次扩张.

可以用已知的单代数扩张的存在性得到一般的关于根域存在性命题, 如下所述:

定理 1 任意域上的任意多项式都有一个根域.

对于一次多项式, 其根域刚好是基域F; 因此我们可以对$f(x)$的次数n用归纳法. 假定这个定理对所有域F和所有$n-1$次多项式都成立, 并设$p(x)$是已知多项式$f(x)$的在F上一个不可约因子. 根据§14.3的定理5, 存在一个由$p(x)$的根u生成的单扩张$K=F(u)$. 在K上, $f(x)$有根u, 因而有因子$x-u$, 所以$f(x)=(x-u)g(x)$, 商$g(x)$是K上$n-1$次多项式, 根据归纳法假定, 由$g(x)$的$n-1$个根生成K上的一个根域N. 这个域N也是$f(x)$的根域.

我们将在下一节(定理2)证明, 给定基域F上的已知多项式$f(x)$的所有根域是同构的, 所以说它是$f(x)$在F上的根域是合理的.

附录 定理1可以用来(纯代数地)构造任意有限域或可数域 F 的代数完全扩张，如下所述．域 F 上的 n 次多项式的个数是有限的或可数的，如果 F 是可数的，那么多项式个数是 $\mathbf{d}^{n+1}=\mathbf{d}$ ($\mathbf{d}=$ 无限可数)．因此 F 上所有多项式的个数是可数的(参看§12.2习题14)，于是我们可以把这些多项式排成序列 $p_1(x)$，$p_2(x)$，$p_3(x)$，$\cdots$．

现在设 F_1 是 $p_1(x)$ 在 F 上的根域，F_2 是 $p_2(x)$ 在 F_1 上的根域，$\cdots$；一般地，设 F_n 是 $p_n(x)$ 在 F_{n-1} 上的根域．最后，设 F^* 是由出现在其中一个 F_n 中，因而出现在 F_n 的所有后继中的所有元素组成的集合．如果 a 和 b 是 F^* 的任意两个元素，那么它们一定都在某一个 F_n 中，因而都在 F_n 的所有后继中．所以 $a+b$，ab 和 $\frac{a}{b}$ $(b\neq 0)$ 在 F_n 和它的所有后继中也一定有相同的值，这就表明 F^* 是一个域．

为了证明 F^* 是代数完全域，我们设 $g(x)$ 是 F^* 上任意多项式，$g(x)$ 的全部系数都在某一个 F_n 中，因而这些系数在 F 上是代数的．那么利用§14.5定理9，我们可以求出 $g(x)$ 的一个非零倍式 $h(x)$，其系数在 F 中(见下面的习题5)．但是对 $h(x)$，有一个适当的 F_{m-1}，$h(x)$ 在 F_{m-1} 上的根域是 F_m，$h(x)$ 在 F_m 中一定可以分解成线性因子，因此 $h(x)$ 的因子 $g(x)$ 在 F_m 中同样也可分解成线性因子．所以 $g(x)$ 在较大的域 F^* 上也可以分解成线性因子，因此 F^* 是代数完全域．进一步有，F^* 的每个元素在 F 上是代数的。

用一般的良序集合和所谓超限归纳法来代替序列，可把上述推理过程加以修改①，以便应用到任意域 F 上．这种修改建立了

① 详细证明见 B. L. van dev Waerden, *Moderne Algebra*, Part I, Berlin, 1930.（中译本：B. L. 范德瓦尔登，代数学 I，丁石孙等译，科学出版社，1965.）

下面关于代数基本定理的重要的部分的推广：任意域F都有一个代数完全扩张。

§15.2 唯一性定理

我们现在来证明定理1所述的根域的唯一性(精确到同构).

定理2 域F上已知多项式$f(x)$的任意两个根域N和N'同构. N到N'的同构可以如此选择，使得F的元素保持固定.

证明 "根域是唯一的"这个断言实质上是"同一个不可约多项式的两个不同根生成同构的单代数扩张"这一事实的直接推论(§14.3定理6). 特别是，不可约多项式$p(x)$的两个根域$N=F(u_1, \cdots, u_n)$和$N'=F(u'_1, \cdots, u'_n)$分别包含由$p(x)$的根u_1和u'_1生成的同构单扩张$F(u_1)$和$F(u'_1)$. 因此存在$F(u_1)$到$F(u'_1)$的同构T. 现在只需要把这个同构适当地扩张到整个根域上. 这种扩张的基本方法由下面引理给出.

引理1 如果域F和F'之间的同构S把不可约多项式$p(x)$的系数映射到F'上多项式$p'(x)$相应的系数，并设$F(u)$和$F'(u')$分别是由这两个多项式的根u和u'生成的单扩张，那么S可以扩张成$F(u)$到$F'(u')$的同构S^*，在这个同构之下，$uS^*=u'$.

证明 恰好同§14.3定理6的讨论一样，我们所需要的同构S^*由公式

$$(a_0+a_1u+\cdots+a_{n-1}u^{n-1})S^* \\ =a_0S+(a_1S)u'+\cdots+(a_{n-1}S)(u')^{n-1} \tag{1}$$

明显地给出，其中所有$a_i\in F$, n是u在F上的次数.

引理2 如果F到F'的同构S把多项式$f(x)$映射到$f'(x)$，并设$N\supset F$和$N'\supset F'$分别是$f(x)$和$f'(x)$的根域，那么同构S可以扩张成N到N'的同构.

通过对次数$m=[N:F]$用归纳法可以证明这个引理. 当$m=$

1时，这是显然的，因为这时S已经扩张到N；因此取$m>1$，并且假定引理对于某域F上次数小于m的所有根域N都正确。因为$m>1$，$f(x)$的根不全都在F中，所以$f(x)$中至少有一个不可约因子$p(x)$，它的次数为$d>1$。设u是$p(x)$在N中的根，而在给定的同构S之下，$p'(x)$是$f'(x)$的对应于$p(x)$的因子。那么根域N'包含$p'(x)$的根u'，并根据引理1，给定的同构S可以扩张成同构S^*，满足

$$
\begin{aligned}
&uS^*=u',\quad [F(u)]S^*=F'(u'),\\
&p(u)=0,\quad p'(u')=0.
\end{aligned}
\tag{2}
$$

因为N是在F上添加$f(x)$的全部根而生成，N一定是在较大的域$F(u)$上添加这些根而生成，所以N是$f(x)$在$F(u)$上的根域，其次数是$\frac{m}{d}$。根据同样的理由，N'是$f'(x)$在$F'(u')$上的根域。因为$\frac{m}{d}<m$，所以根据引理的归纳假定可以断言，(2)表示的同构S^*可以从$F(u)$扩张到N。这就证明了引理2。

两个根域N和N'都是同一个基域F的扩张，并且S是F到自身上的恒等映射，这种情形下，引理2表明N与N'同构，因此也就证明了定理2。

习　　题

1．求下列多项式在$\mathbf{Q}$上的根域的次数：

(a) $x^3-x^2-x-2=0$,

(b) $x^3-2=0$,

(c) $x^4-7=0$,

(d) $(x^2-2)(x^2-5)=0$.

2．证明：域F上n次多项式的根域在F上的次数至多是$n!$.

3．(a) 证明：如果ζ是n次本原单位根，那么$\mathbf{Q}(\zeta)$是$x^n-1=0$在$\mathbf{Q}$上的根域.

(b) 当 $n=3, 4, 5, 6$ 时，计算根域的次数.

4. 证明：特征为 p 的任意代数完全域包含一个子域与 §15.1 附录中所构造的域同构.

*5. 设 $g(x)=a_0+a_1x+\cdots+a_nx^n$ 的系数在域 F 上是代数的，证明：$g(x)$ 是系数在 F 中的某非零多项式 $h(x)$ 的因子.（提示：构造 $g(x)$ 在 $F(a_0, \cdots, a_n)$ 上的根域；在这个根域上把 $g(x)$ 分解成线性因子 $x-u_i$；u_i 在 F 上是代数的，并满足不可约多项式方程 $h_i(x)$；置 $h(x)=\Pi h_i(x)$.）

6. 设 $p\in\mathbf{Q}[x]$ 是任意有理系数首一多项式，并设 $z_1, \cdots, z_n$ 是它的复根. 证明：$\mathbf{Q}(z_1, \cdots, z_n)$ 是 p 在 $\mathbf{Q}$ 上的根域.

§15.3 有限域

系统地运用根域的性质，我们可以得到所有有限域（含有有限个元素的域）的一个完整论述. 因为特征为∞的域总包含一个与有理数域同构的无限子域（§13.8 定理 14），所以每个有限域具有一个素特征 p. 不失一般性，我们可以假定 F 包含模 p 整数域 $\mathbf{Z}_p$（见 §13.7 定理 12 的推论）. 那么有限域 F 是 $\mathbf{Z}_p$ 的有限扩张，所以在 $\mathbf{Z}_p$ 上有一组基底 $u_1, \cdots, u_n$. F 中每个元素可唯一地表示成线性组合 $\sum a_iu_i$. 这里每个系数恰好可按 p 种方法在 $\mathbf{Z}_p$ 中取值，所以 F 中总共有 p^n 个元素. 这就证明了

定理 3 有限域中元素的个数 q 是它的特征的幂 p^n.

在含有 $q=p^n$ 个元素的有限域 F 中，全体非零元素构成 $q-1$ 阶乘法群. 那么这个群中每个元素的阶是 $q-1$ 的一个因子. 所以 F 的每个元素满足方程 $x^{q-1}=1$. 因此 F 的所有元素 $a_1, a_2, \cdots, a_q$（包含着零）满足方程

$$x^q-x=0,\quad q=p^n. \tag{3}$$

因此乘积 $(x-a_1)(x-a_2)\cdots(x-a_q)$ 是 x^q-x 的因子，乘积中的这些因子是互素多项式，每一个都整除 x^q-x. 因为这个乘积同 x^q-x 一样，也是 q 次的首一多项式，所以我们得到

$$x^q-x=(x-a_1)(x-a_2)\cdots(x-a_q). \tag{4}$$

因此F是x^q-x在$\mathbf{Z}_p$上的根域. 元素个数相同的任意其他有限域F'是同一个方程的根域,因此根据根域的唯一性定理(定理2),它与F同构. 这个推理证明了

定理4 元素个数相同的任意两个有限域同构.

下面考虑这样一个问题: 实际上存在哪些有限域? 为了列出一个有限域,我们自然要构造多项式x^q-x在$\mathbf{Z}_p$上的根域N. 我们现在证明,所要求的根域恰好由这个多项式的全部根组成.

引理 多项式 x^q-x 在它的根域N中有q个不同的线性因子.

用反证法来证明. 如果x^q-x有一个重因子$x-u$, 我们可以写成$x^q-x=(x-u)^2g(x)$. 比较它们的形式导数(§3.1 习题7),我们将有

$$(x^q-x)'=qx^{q-1}-1=-1,$$

$$[(x-u)^2g(x)]'=(x-u)[2g(x)+(x-u)g'(x)],$$

因此$x-u$是-1的因子,得出矛盾. 这就证明了引理.

另一方面,x^q-x的根$u_1, \cdots, u_q$中任意两个根之和是一个根,这是因为, 在特征为p的任意域中有$(a\pm b)^p=a^p\pm b^p$, 所以如果$a^{p^n}=a$和$b^{p^n}=b$,那么

$$(a\pm b)^{p^n}=a^{p^n}\pm b^{p^n}=a\pm b.$$

两个根之积ab也是一个根,这因为$(ab)^{p^n}=a^{p^n}b^{p^n}=ab$. 还有对两个根之商, 类似的结果成立. 所以$x^q-x$的所有$q$个根组成的集合是根域$N$的一个子域;因为这个子域包含所有的根,所以它实际上一定是整个根域N. 这就意味着, 我们已经构造出含有q个元素的域, 因此有

定理5 对任意素数p和任意正整数n, 存在一个含有$p^n=q$个元素的有限域: $x^q-x=0$在$\mathbf{Z}_p$上的根域.

根据定理 4 和定理 5，存在一个且本质上只存在一个含有 p^n 个元素的域．这个域有时称为伽罗瓦域 $GF(p^n)$．这个域的乘法群的结构可以完整地描述如下．

定理 6 *在任意有限域 F 中，所有非零元素组成的乘法群是循环群．*

证明 F 中每个非零元素是 $q-1$ 次单位根，即它满足方程 $x^{q-1}=1$，这里 q 是 F 中元素的个数．为了证明这个群是循环群，我们必须求出 F 中 $q-1$ 次本原单位根，它的低于 $q-1$ 次的幂都不等于 1；那么本原单位根的所有次幂跑遍整个群．为此，把 $q-1$ 写成不同素数幂的乘积

$$q-1=p_1^{e_1}p_2^{e_2}\cdots p_r^{e_r} \qquad (0<p_1<p_2<\cdots<p_r).$$

对每个 $P=p_i$，有 $P^e|(q-1)$，所以 $x^{P^e}=1$ 的全部根都是 $x^{q-1}=1$ 的根，因此它们都在 F 中．这个方程 $x^{P^e}=1$ 的所有 P^e 个不同的根中，恰有 P^{e-1} 个根满足方程 $x^{P^{e-1}}=1$，所以 F 至少包含 $x^{P^e}=1$ 的一个根 $c=c_i$，它不满足 $x^{P^{e-1}}=1$．于是这个元素 c_i 在 F 的乘法群中的阶数为 $p_i^{e_i}$．乘积 $c_1c_2\cdots c_r$ 是 $q-1$ 阶元素(参看下面的习题 8)，这正是所要求的．

定理 7 *每个特征为 p 的有限域有一个自同构 $a\longleftrightarrow a^p$．*

证明 从特征为 p 的域的一般讨论中我们知道，对应 $a\longmapsto a^p$ 把 F 同构地映射到它元素的 p 次幂组成的集合(§13.7 定理 13)．因为这个对应是一一的，所以 q 个元素 a 恰好给出 q 个 p 次幂，那么它们一定包含整个域 F．因此 $a\longmapsto a^p$ 把 F 映射到整个 F 上．

推论 *在特征为 p 的有限域中，每个元素有 p 次根．*

有限域的另外一些性质将在习题中叙述．

习　　题

1．证明：对每个正整数 n，都存在一个 $\mathbf{Z}_p$ 上 n 次不可约多项式．

2. 证明：包含 $\mathbf{Z}_p$ 的每个有限域是 $\mathbf{Z}_p$ 的单扩张.

3. 证明：有限域的每个有限扩张是单扩张.

4. (a) 用次数证明：$GF(p^n)$ 的任意子域有 p^m 个元素，这里 $m|n$.

(b) 证明：如果 $m|n$，那么 $(p^m-1)|(p^n-1)$.

(c) 用(b)证明：如果 $m|n$，那么 $GF(p^n)$ 有一个含有 p^m 个元素的子域.

5. 证明：p^n 阶伽罗瓦域的所有子域组成的格与 n 的所有正因子组成的格同构.

6. 证明：在 $GF(p^n)$ 中自同构 $a\longleftrightarrow a^p$ 的阶为 n.

7. 证明：如果 m 与 F 的特征 p 互素，那么在 F 上存在一个 m 次本原单位根.（提示：应用证明定理 6 时用过的方法．这个方法可以应用到特征为 ∞ 的域上吗？）

8. 证明：在阿贝耳群中，每个元素 c_i 的阶是不同素数的幂 $p_i^{e_i}$，这些元素的乘积 $c_1c_2\cdots c_r$ 的阶恰好是 $p_1^{e_1}\cdots p_r^{e_r}=h$.（提示：指出这个乘积的阶数可整除 h，但对每个 i，它不能整除 $\frac{h}{p_i}$.）

9. (a) 用基本原理证明：模 p 的非零整数(在 $\mathbf{Z}_p$ 中)组成的乘法群是循环群.

(b) 设 ζ 是有理数域 $\mathbf{Q}$ 上 p 次本原单位根，利用(a)来证明：$\mathbf{Q}(\zeta)$ 在 $\mathbf{Q}$ 上的伽罗瓦群是 $p-1$ 阶循环群.

10. (a) 证明：在阶数为 $q=p^n$ 的任意有限域中，由完全平方组成的集合 S 的基数至少是 $\frac{q+1}{2}$.

(b) 验证：对任意 $a\in S$，集合 $S\cap(a-S)$ 不可能是空集.

(c) 推出每个元素是两个平方之和.

§15.4 伽罗瓦群

群不仅可以用来表示几何图形的对称，而且还可以表示代数系统的对称．例如，复数域 $\mathbf{C}$ 相对于实数域而言有两种对称：一个是恒等，另一个是同构 $a+bi\longleftrightarrow a-bi$，这个同构把每个数映射到它的复共轭．这种一个域到自身的同构称为自同构．一般

地，域K的自同构T是集合K到它自身的双射，并保持和与积，即对K中所有a和b，有

$$(a+b)T=aT+bT,\ (ab)T=(aT)(bT). \tag{5}$$

两个自同构S和T的乘积ST也是一个自同构，并且自同构的逆仍然是自同构. 因此

定理 8 *域K的所有自同构组成的集合在乘积之下构成一个群.*

设K是F的扩张，并考虑这样一些自同构T，它对F中每个元素a，有$aT=a$. 也就是说这些自同构保持F中任何元素都不变. 在K的整个自同构群中，它们构成一个子群，称为K在F上的自同构群. 例如，$\mathbf{C}$在$\mathbf{R}$上的自同构群由两个自同构$a+bi\longleftrightarrow a+bi$和$a+bi\longleftrightarrow a-bi$组成.

定义 *域K在子域F上的自同构群是由保持F的元素不变的K的那些自同构组成的群.*

最重要的特例是代数数域在有理数域$\mathbf{Q}$上的自同构群，但是在我们考虑具体例子之前，先让我们确定代数数在自同构之下可能的象.

定理 9 *域F的有限扩张K的任意自同构T把K的每个元素u映射到u在F上的共轭元素uT.*

这个定理断言，u和它的象uT都满足F上同一个不可约方程. 为证明这一点，设给定的元素u在F上是代数的，它满足一个系数在F中的首一不可约多项式方程$p(x)=x^n+b_{n-1}x^{n-1}+\cdots+b_0=0$. 根据公式(5)，自同构$T$保持所有有理关系，并保持每个$b_i$固定，因此由$p(u)=0$得到

$$\begin{aligned}&(u^n+b_{n-1}u^{n-1}+\cdots+b_0)T\\&=(uT)^n+b_{n-1}(uT)^{n-1}+\cdots+b_1(uT)+b_0=0.\end{aligned}$$

这个方程表明，uT也是$p(x)$的根，因此uT是u的共轭.

例 1 考虑有理数域上的四次域①，$K=\mathbf{Q}(\sqrt{2},i)$，它是由$\sqrt{2}$和$i=\sqrt{-1}$生成的. 整个域K是中间域$F=\mathbf{Q}(i)$上的二次扩张，它是由$x^2=2$的两个共轭根$\pm\sqrt{2}$中的任意一个生成的. 根据§ 14.3 定理 6，存在K的自同构S，把$\sqrt{2}$映射到$-\sqrt{2}$，并保持$\mathbf{Q}(i)$中元素固定. 也就是说，共轭根$\sqrt{2}$和$-\sqrt{2}$在代数上没有什么差别. S作用到K的任意元素u上，是

$$(a+b\sqrt{2}+ci+d\sqrt{2}i)S=a-b\sqrt{2}+ci-d\sqrt{2}i, \qquad (6)$$

这里我们通过基底$1,\sqrt{2},i,\sqrt{2}i$把K的每个元素写出来(参看§ 14.5). 通过类似的论证，存在一个自同构T，它保持$\mathbf{Q}(\sqrt{2})$的元素固定，并把i映射到$-i$. 那么有

$$(a+b\sqrt{2}+ci+d\sqrt{2}i)T=a+b\sqrt{2}-ci-d\sqrt{2}i, \qquad (7)$$

所以T就是把每个数映射到它的复共轭. 乘积ST是K的第三个自同构. 这些自同构作用到$\sqrt{2}$和i上的效果可以列表如下:

$$S:\ \begin{cases}\sqrt{2}\longmapsto-\sqrt{2}\\ i\longmapsto i,\end{cases} \qquad T:\ \begin{cases}\sqrt{2}\longmapsto\sqrt{2}\\ i\longmapsto -i,\end{cases}$$

$$ST:\ \begin{cases}\sqrt{2}\longmapsto-\sqrt{2}\\ i\longmapsto -i,\end{cases} \qquad I:\ \begin{cases}\sqrt{2}\longmapsto\sqrt{2}\\ i\longmapsto i.\end{cases}$$

我们断言，I,S,T和ST是K在$\mathbf{Q}$上的仅有的自同构. 根据定理 9，任意其他的自同构U一定把$\sqrt{2}$映射到共轭数$\pm\sqrt{2}$，把i映射到共轭数$\pm i$. 恰好有四种可能性，就是上面表中列出的I，S，T和ST. 因此U作用到生成元$\sqrt{2}$和i上的效果一定同这四种自同构之一是一致的，因此它作用到整个域上的效果也是一致的. 于是$U=I$或S,T,ST.

这些自同构的乘法表可以直接由上面列出的作用到$\sqrt{2}$和i的表求出. 它是

$$S^2=I,T^2=I,ST=TS. \qquad (8)$$

① 同§ 14.5 中一样，我们可以把这个域看作x^4-2x^2+9的根域.

这完全象四群的乘法表(见§6.7)，于是我们得出结论：$\mathbf{Q}(\sqrt{2}, i)$的自同构群与四群$\{I, S, T, ST\}$同构.

定义 如果$N=F(u_1, \cdots, u_n)$是多项式$f(x)=(x-u_1)\cdots(x-u_n)$的根域，那么$N$在$F$上的自同构群称为方程$f(x)=0$的伽罗瓦群，或者称为域$N$在$F$上的伽罗瓦群.

为了明显地描述特殊伽罗瓦群的自同构T，我们进行如下讨论．设N是$f(x)$在F上的根域，那么T把$f(x)$的根映射到$f(x)$的根(定理9)，并把不同的根映射到不同的根．因此T的作用相当于对$f(x)$所有不同的根$u_1, \cdots, u_k$作一个置换ϕ，所以

$$u_1T=u_{1\phi}, \cdots, u_kT=u_{k\phi}, k\leqslant n. \tag{9}$$

另一方面，根域中每个元素w可以表示成多项式$w=h(u_1, \cdots, u_k)$，其系数在F中．因为T保持这些系数固定，所以由T的性质(9)得到

$$\begin{aligned}[h(u_1, \cdots, u_k)]T&=h(u_1T, \cdots, u_kT)\\&=h(u_{1\phi}, \cdots, u_{k\phi}).\end{aligned}$$

这个公式表明，T作用到w上的效果完全由T作用到根上的效果确定，或者说，T由置换(9)唯一确定．因为两个置换的乘积是通过相继作用相应的自同构而得到，所以全体形为(9)的置换构成一个群，与自同构群同构．置换(9)只包含这样一些置换：它保持全体根之间的所有多项式恒等式不变，所以它对应于自同构．如此建立的结果可以概述如下：

定理 10 设$f(x)$是F上任意n次多项式，它在根域$N=F(u_1, \cdots, u_k)$中恰有k个不同的根$u_1, \cdots, u_k$，那么$f(x)$的伽罗瓦群G的每个自同构T诱导出一个作用在$f(x)$的全体不同根上的置换$u_i \longleftrightarrow u_iT$，而且$T$由这个置换完全确定.

推论 1 任意多项式的伽罗瓦群与它的根的置换群同构.

推论 2 n次多项式的伽罗瓦群的阶可整除$n!$.

例2 根据爱森斯坦定理，方程 $x^4-3=0$ 在域 $\mathbf{Q}$ 上是不可约的，并有四个根 $r, ir, -r, -ir$，这里 $i=\sqrt{-1}$，$r=\sqrt[4]{3}$ 是 3 的正实四次根．根域 $N=\mathbf{Q}(r, ir, -r, -ir)$ 可以看作是由 r，i 生成的，即 $N=\mathbf{Q}(r, i)$．因为 r 在 $\mathbf{Q}$ 上是四次的，i 是复数，因此 i 在实域 $\mathbf{Q}(r)$ 上是二次的，所以整个根域 N 在 $\mathbf{Q}$ 上的次数是 8．根据 § 14.5 定理 9，这个扩张有 8 个元素 $1, r, r^2, r^3, i, ir, ir^2, ir^3$ 构成的一组基底．因为 N 中每个元素可以表示成这些基元素的线性组合，其系数为有理数，所以自同构 T 的作用只要知道了 rT 和 iT，就马上完全确定了．

可以很容易地构造出 N 的几个自同构．因为 N 是实域 $\mathbf{Q}(r)$ 上的二次扩张，所以它有一个把 N 中每个数映射到它的复共轭的自同构，因此 $rT=r, iT=-i$．另一方面，N 是子域 $\mathbf{Q}(i)$ 的四次扩张，它由元素 r 生成．根据 § 14.3 定理 6，N 有一个自同构 S，把 r 映射到它的共轭 ir，所以 $rS=ir, iS=i$．由此得到，S^2 是一个自同构，它满足 $rS^2=i^2r, iS^2=i$，而 $rS^3=-ir$，$iS^3=i$．把 S 和 T 进一步组合，我们得到 N 的八个自同构，它们作用到生成元 r 和 i 上的效果如下：

	I	S	S^2	S^3	T	TS	TS^2	TS^3
把 r 映射到	r	ir	$-r$	$-ir$	r	ir	$-r$	$-ir$
把 i 映射到	i	i	i	i	$-i$	$-i$	$-i$	$-i$

我们还可以计算 $TS^3=ST, S^4=T^2=I$，所以这八个自同构构成一个群，它与正方形对称群（§ 6.4）同构．这些自同构组成整个伽罗瓦群，这因为任意自同构一定把 i 映射到它的共轭 $\pm i$ 中的一个，把 r 映射到共轭 $\pm r$ 或 $\pm ir$．上面这个表包括这些作用的所有八种可能的组合．

群论的许多概念可以应用到这样的伽罗瓦群 G 上．例如，G

包含着由 S 生成的子群 $H=[I, S, S^2, S^3]$ 和由 S^2 生成的较小的子群 $L=[I, S^2]$. 子群 H 的每个自同构保持 i 不变，因而保持子域 $\mathbf{Q}(i)$ 中每个元素不变. 较小的子群 L 由这样的自同构组成，这些自同构保持较大的子域 $\mathbf{Q}(i, r^2)$ 中每个元素不变. 在这种意义下，下降的子群序列 $G\supset H\supset L\supset I$ 对应着上升的子域序列 $\mathbf{Q}\subset\mathbf{Q}(i)\subset\mathbf{Q}(i, \sqrt{3})\subset\mathbf{Q}(i, r)$. 于是，这一上升的子域序列给出求解已知方程的方法，它是通过逐次添加较简单的方程 $x^2=-1$, $y^2=3$, $z^2=\sqrt{3}$ 的根来实现的. 这个例子说明了伽罗瓦群的子群对于求方程的根式解所起的作用.

自然要考虑伽罗瓦群的同态. 上述群 G 的每个自同构 U 把 i 映射到 $\pm i$，因而把域 $\mathbf{Q}(i)$ 中每个元素映射到同一个域中的某个元素. 这就意味着，U 诱导出 $\mathbf{Q}(i)$ 的一个自同构 U^*，这里 U^* 是对 $\mathbf{Q}(i)$ 中的元素 w 通过等式 $wU^*=wU$ 来定义的. 对应 $U\longmapsto U^*$ 是一个同态，它把 N 的所有自同构 U 组成的群 G 映射到 $\mathbf{Q}(i)$ 的自同构组成的群 G^* 上. 但是 G^* 只有两个元素：恒等映射 I^* 和对换 i 和 $-i$ 的自同构. 进一步，$U^*=I^*$ 当且仅当 U 保持 $\mathbf{Q}(i)$ 的每个元素不变，也就是当且仅当 U 在子群 $H=[I, S, S^2, S^3]$ 中. 因此 $U\longmapsto U^*$ 是 G 的满同态，它的核是 H，因此群 G^* 与商群 G/H 同构.

习　题

1. 对 $\mathbf{Q}(i, r)$ 的所有子域组成的系统画一个格图.

2. 通过证明 x^4-3 没有系数在 $\mathbf{Q}$ 中的线性因子和二次因子来证明：x^4-3 在 $\mathbf{Q}$ 上是不可约的.

3. 把 x^4-3 的伽罗瓦群的每个自同构表示成它的根的置换.

4. (a) 证明：x^4-3 在 $\mathbf{Q}(i)$ 上是不可约的.

   (b) 描述 x^4-3 在 $\mathbf{Q}(i)$ 上的伽罗瓦群.

5. 用基本原理证明：x^4-3 根的置换：$r\mapsto ir$, $ir\mapsto -ir$, $-ir\mapsto -r$,

$-r \mapsto r$，不可能对应于一个自同构.

6. 设 $F=\mathbf{Q}(\omega)$ 是由三次复单位根 ω 生成的域. 讨论 x^3-2 在 F 上的伽罗瓦群，包括：确定这个根域的次数，用纯群论的语言来描述伽罗瓦群，把每个自同构表示成置换.

7. 按照习题 6 的内容讨论 x^5-7 在 $\mathbf{Q}(\zeta)$ 上的伽罗瓦群，这里 ζ 是五次本原单位根.

8. 证明：有限域的伽罗瓦群是循环群.

9. 证明：如果 ζ 是 n 次本原单位根，那么 $\mathbf{Q}(\zeta)$ 的伽罗瓦群是阿贝耳群. (提示：任意自同构有这样的形式 $\zeta \mapsto \zeta^e$.)

10. (a) 证明：如果 K 是 $\mathbf{Q}$ 的扩张，那么 K 的每个自同构保持 $\mathbf{Q}$ 的每个元素不变.

(b) 叙述并证明对于特征为 p 的域有类似的结果.

§15.5 可分多项式与不可分多项式

由于存在所谓的不可分不可约多项式或不可分元素——即这些元素是 n 次代数的，但它的共轭元素的个数小于 n——伽罗瓦群的一般讨论就变得复杂了. 对某些特征为 p 的域，这种复杂化就出现了，这可以用简单的例子加以说明.

设 $K=\mathbf{Z}_p(u)$ 表示模 p 整数的域 $\mathbf{Z}_p$ 的单超越扩张，并设 F 表示由 $u^p=t$ 生成的 K 的子域 $\mathbf{Z}_p(u^p)$. 于是，F 是由 $\mathbf{Z}_p$ 上的超越元素 t 的所有有理形式组成. 原来的元素 u 满足 F 上的一个多项式方程 $f(x)=x^p-t=0$. 这个多项式 $f(x)$ 在 $F=\mathbf{Z}_p(t)$ 上实际上是不可约的，这是因为如果 f 在 $\mathbf{Z}_p(t)$ 上可约，根据高斯引理(§3.9)，在 t 的多项式整环 $\mathbf{Z}_p[t]$ 上，f 是可约的，但是，由于 $f(x)=x^p-t$ 对于 t 来说是线性的，所以这样的因式分解 $f(x)=g(x, t)h(x, t)$ 是不可能的. 因此 $f(x)$ 的根 u 在 F 上的次数是 p. 但是 $f(x)$ 在 K 上有因式分解

$$f(x)=x^p-u^p=(x-u)^p. \tag{10}$$

因此它只有一个根 u，并且 u（虽然它的次数 $p>1$）除了它本身之外没有其他共轭元素.

我们可以用下面的术语来描述上述情况.

定义 域 F 上的一个 n 次多项式 $f(x)$，如果它在某个根域 $N\supseteq F$ 中有 n 个不同的根，那么称它在域 F 上是可分的；否则，称 $f(x)$ 是不可分的. 如果有限扩张 $K\supseteq F$ 中每个元素在 F 上都满足一个可分多项式方程，那么称 K 在 F 上是可分的.

容易检验给定的多项式 $f(x)=a_0+a_1x+\cdots+a_nx^n$ 是否是可分的. 这也就是，首先根据公式（参看 §3.1 习题 7）

$$f'(x)=a_1+(2\times a_2)x+\cdots+(n\times a_n)x^{n-1} \tag{11}$$

来定义 $f(x)$ 的形式导数 $f'(x)$，这里 $n\times a_n$ 表示 a_n 的 n 次自然倍数（见 §13.7）. 如果这些系数都在实数域中，那么这种导数与微积分学中建立的普通导数是一致的. 从形式导数的定义(11)出发，不用任何极限概念，我们可以推导出很多微分法则，例如

$$(f+g)'=f'+g',\quad (fg)'=fg'+f'g,\quad (f^m)'=mf^{m-1}f',$$

等等.

现在在任意根域 N 上把 $f(x)$ 分解成不同线性因子的幂

$$f(x)=c(x-u_1)^{e_1}\cdots(x-u_k)^{e_k}\quad (c\neq 0). \tag{12}$$

把(12)的两边形式地微分，我们看到，$f'(x)$ 是 $ce_1(x-u_1)^{e_1-1}(x-u_2)^{e_2}\cdots(x-u_k)^{e_k}$ 与 $k-1$ 项每项都包含 $(x-u_1)^{e_1}$ 作为因子的和. 因此当 $e_1>1$ 时，$x-u_1$ 可整除 $f'(x)$，而当 $e_1=1$ 时，$x-u_1$ 就不能整除 $f'(x)$. 对 $e_2,\cdots,e_k$ 重复上述推理，我们得到 $f(x)$ 和 $f'(x)$ 有公因子，除非 $e_1=e_2=\cdots=e_k=1$，即除非 $f(x)$ 是可分的；因此 $f(x)$ 在 N 上是可分的当且仅当 $f(x)$ 和它的形式导数 $f'(x)$ 是互素的.

$f(x)$ 和 $f'(x)$ 的最大公因式可象第三章中那样用 $F[x]$ 中的欧几里得算法直接计算出来；当 F 扩张到较大的域上，它们的最大公因式并不改变. 于是我们得到

定理 11 设 $f(x)$ 是域 F 上任意多项式，用欧几里得算法计算 $f(x)$ 和它的形式导数 $f'(x)$ 的最大公因式 $d(x)$（首一多项式）。如果 $d(x)=1$，那么 $f(x)$ 是可分的；否则 $f(x)$ 是不可分的。

如果 $f(x)$ 是不可约的，那么除非 $f(x)$ 整除 $g(x)$，总有 g. c. d. $(f(x), g(x))=1$，并且 $f(x)$ 不能整除任意较低次的非零多项式。因此得到

推论 1 任意不可约多项式是可分的，除非它的形式导数是零。

推论 2 特征为∞的域上的任意不可约多项式是可分的。

这是因为，当 $n>0, a_n\neq 0$ 时，$f'(x)=n\times a_n x^{n-1}+\cdots\neq 0$。

进一步推论是：如果 F 的特征是∞，那么任意 n 次不可约多项式的根域恰恰包含 $f(x)$ 的 n 个不同的共轭根。而且，在特征为∞的域上的任意代数数满足一个不可约的因而也是可分的方程，所以特征为∞的域的任意代数扩张，在上述意义之下是可分扩张。

推论 2 的结果对于素特征的域是不成立的。例如，在本节开始所提到的不可约多项式 $f(x)=x^p-t$ 有形式导数 $p\times x^{p-1}=0$。

习　　题

1. 不用定理 11 证明：$\mathbf{Q}$ 上二次不可约多项式的根是不同的。

2. 设 $f(x)$ 是有理系数多项式，而 $d(x)$ 是 $f(x)$ 和 $f'(x)$ 的最大公因式。证明：$\dfrac{f(x)}{d(x)}$ 是与 $f(x)$ 具有相同根的多项式，但它没有重根。

3. (a) 证明：如果 $f'(x)=0$，那么 $f(x)$ 在任意域 F 上是不可分的。

*(b) 证明：如果在 $\mathbf{Z}_p$ 上 $f'(x)=0$，那么对某个 $g(x)$，有

$$f(x)=[g(x)]^p.$$

4. 证明：x^3-2u 在 $\mathbf{Z}_3(u)$ 上是不可分的。证明：它的根域的伽罗瓦群是由一个单位元素组成。

5. 用定理 11 证明：当 $q=p^n$ 时，x^q-x 在 $\mathbf{Z}_p$ 上是可分的。

6. (a) 证明：如果 $f(x)$ 是在特征为 p 的域 F 上的满足 $f'(x)=0$ 的一

个多项式，那么 $f(x)$ 可以写成形式 $a_0+a_1x^p+\cdots+a_nx^{np}$.

(b) 证明：如果 F 是有限域，那么对某适当的 $g(x)$，有

$$f(x)=[g(x)]^p.$$

(c) 用(b)证明：有限域上的每个不可约多项式是可分的.

§15.6 伽罗瓦群的性质

可分多项式的根域和伽罗瓦群有两个极好的性质，现在我们把它们叙述成定理.

定理 12 *域 F 上可分多项式的伽罗瓦群的阶恰好等于它的根域的次数 $[N:F]$.*

在 §15.4 的第二个例子中我们已经看到，多项式 $x^4-3=0$ 的根域就是这种情形.

定理 13 *在可分多项式的根域 $N\supset F$ 中，在 N 的(在 F 上的)伽罗瓦群的每个自同构之下保持不变的元素恰恰就是 F 的元素.*

这个定理告诉我们一些关于伽罗瓦群正面的信息，因为它断言，对 N 中每个不在 F 中的元素 a，在 G 中有一个自同构 T，使得 $aT\neq a$.

为证明定理 12，参考 §15.2 的引理 2，它与域之间同构的可扩张性有关. 注意，在这个引理中(不象 §15.5 中那样)，$f'(x)$ 并不表示 $f(x)$ 的导数.

引理 *如果 §15.2 引理 2 中的多项式 $f(x)$ 是可分的，那么 S 可以按照 $m=[N:F]$ 种不同的方法扩张到 N.*

这个结果可以通过对 m 用数学归纳法来证明. 域 F 到 F' 的已知同构 S 的任意扩张 S^*，按照(2)式把根 u 映射到 $p'(x)$ 的某个根 u'，因此 S 的每种可能的扩张可以通过上述一种构造而得到. 因为 $f(x)$ 是可分的，所以它的 d 次因子 $p(x)$. 恰有 d 个不同的根 u. u' 的 d 种选择恰好给出(2)中 S^* 的 d 种选择. 根据归纳法假

设，每个这样的 S^* 可以按照 $\frac{m}{d}=[N:F(u)]$ 种不同的方法扩张到 N，所以总共有 $d\left(\frac{m}{d}\right)=m$ 种扩张，如断言所述．

如果 $f(x)=f'(x)$ 是 m 次可分的，在 §15.2 引理 2 中我们令 $N=N'$，上述引理断言，F 的恒等自同构 I 恰恰可以按照 m 种不同方法扩张到 N 的一个自同构．但是这些自同构组成 N 在 F 上的伽罗瓦群，这就证明了定理 12．

最后，为了证明定理 13，设 G 是可分多项式的根域 N 在域 F 上的伽罗瓦群，而 K 是 N 中所有在 G 的每个自同构之下不变的元素组成的集合．容易证明 K 是一个域，并且 $K\supset F$．因此 G 中每个自同构是 K 的恒等自同构 I 到 N 上的一个扩张．因为 N 是 K 上的根域，所以根据上述引理，只存在 $[N:K]$ 个这样的扩张，而根据定理 12，总共有 $[N:F]$ 个自同构．因此 $[N:K]=[N:F]$．因为 $K\supset F$，所以这就推出 $K=F$，证明了定理 13．

上述关于扩张的引理还有另一个推论，即根域在下述意义下总是“正规的”．

定义 *域 F 的有限扩张 N 称为在 F 上是正规的，是指如果 F 上每个不可约多项式 $p(x)$ 在 N 中有一个根，则它的所有根都在 N 中．*

换句话说，每个在 F 上不可约在 N 中有一个根的多项式 $p(x)$ 在 N 上可以分解成线性因子．

定理 14 *F 的一个有限扩张在 F 上是正规的当且仅当它是 F 上某个多项式的根域．*

证明 如果 N 在 F 上是正规的，那么在 N 中选取任意一个不在 F 中的元素 u，并求出 u 所满足的不可约方程 $p(x)=0$．根据正规性的定义，N 包含 $p(x)$ 的所有根，因此 N 包含 $p(x)$ 的根域 M．如

果N中有不属于M的元素，其中一个元素v满足不可约方程$q(x)=0$，因此M包含在较大的$p(x)q(x)$的根域中，等等．因为N的次数是有限的，所以这样逐次得到的根域中一定有一个是整个域N.

反过来，任意$f(x)$的根域N是正规的．假定有某个多项式$p(x)$在F上是不可约的，它有一个但不是全部根在N中．设w是$p(x)$的位于N中的根，并把另一个不在N中的根w'添加到N上．满足$wT=w'$的对应T是单扩张$F(w)$到$F(w')$的同构．域N是$f(x)$在$F(w)$上的根域；另一方面，$N'=N(w')$是由$f(x)$的根添加到$F(w')$上生成的，因此它是$f(x)$在$F(w')$上的根域．所以根据§15.2引理2，对应T可以扩张成N到N'的同构．因为T保持基域F的元素不变，所以这两个同构的域N和N'在F上的次数一定相同．但是我们已经假定了$N'=N(w')$是N的真扩张，所以N'在F上的次数大于N在F上的次数．由此矛盾故得定理．

如果把这个定理证明的前一半应用到可分扩张上（这个扩张中每个元素都满足一个可分方程），那么所有用到的多项式$p(x)$和$q(x)$都是可分的．这就证明了

推论 F的每个有限正规可分扩张是一个可分多项式的根域．

特别是，有理数域$\mathbf{Q}$的每个有限正规扩张N自然就是可分的（定理11的推论2），因此它是某个可分多项式的根域．因此N在$\mathbf{Q}$上的自同构群的阶恰好就是N在$\mathbf{Q}$上的次数$[N:\mathbf{Q}]$.

伽罗瓦群可以用来研究对称多项式的性质，关于对称多项式已在§6.10中给出定义．

定理15 设$N=F(u_1,\cdots,u_n)$是由n次可分多项式$f(x)$的全部n个根$u_1,\cdots,u_n$生成的域，并设$g(x_1,\cdots,x_n)$是F上n个未定元$x_1,\cdots,x_n$的任意对称多项式形式．那么N的元素$w=g(u_1,\cdots,u_n)$在基域F中．

证明 根据定理 10, N 的伽罗瓦群 G 的任意自同构 T, 它的作用相当于对 $f(x)$ 的根所做的置换 $u_i \longmapsto u_iT$. $g(x_1, \cdots, x_n)$ 的对称性意味着, 对于未定元的任意置换, 它都不变; 因此

$$w \longmapsto wT = g(u_1T, \cdots, u_nT) = g(u_1, \cdots, u_n) = w.$$

因为 w 在任意自同构 T 作用之下都不变, 所以根据定理 13, w 在 F 中.

推论 任意 n 个未定元的对称多项式(F 上的)可以表示成 n 个初等对称函数

$$\begin{aligned} \sigma_1 &= x_1 + x_2 + \cdots + x_n, \\ \sigma_2 &= x_1x_2 + x_1x_3 + \cdots + x_{n-1}x_n, \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ \sigma_n &= x_1x_2\cdots x_n \end{aligned} \tag{13}$$

的有理函数①(F 上的).

为简化公式, 我们只写出 $n=3$ 情形的证明. F 上的初等对称函数 σ_1, σ_2 和 σ_3 生成一个域 $K=F(\sigma_1, \sigma_2, \sigma_3)$. 由原来三个未定元生成的域 $N=F(x_1, x_2, x_3)$ 是 K 的有限扩张, 事实上, N 的生成元 x_1, x_2, x_3 是三次多项式

$$\begin{aligned} f(x) &= (x-x_1)(x-x_2)(x-x_3) \\ &= x^3 - \sigma_1x^2 + \sigma_2x - \sigma_3 \end{aligned}$$

的根, 共中系数原来就是(13)式给出的对称函数. 引进根域 N 在 K 上的伽罗瓦群 G. 根据定理 10, 每个自同构诱导出一个 $x_1, \cdots, x_n$ 的置换, 因此根据定理 15, $x_1, \cdots, x_n$ 的任意对称多项式在基域 K 中. 因为 $K=F(\sigma_1, \sigma_2, \sigma_3)$, 所以由此得出, 这样的对称多项式是 $\sigma_1, \sigma_2, \sigma_3$ 的有理函数.

---

① 参看 §6.10 定理 19, 在那里叙述了一个较强的结果.

习　　题

1. 在定理 15 的推论证明中，证明：$N=K(x_1, x_2, x_3)$在K上的伽罗瓦群恰好是三个字母的对称群.

2. 把$x_1^2+x_2^2+x_3^2$表示成初等对称多项式的有理函数.（也可参看 § 6.10 习题 7 和 8.）

3. (a) 证明：存在域K和子域F，使得K在F上的伽罗瓦群是n次对称群.

(b) 证明：在(a)中，K可以选为实数域的一个子域.（提示：利用n个代数无关的实数.）

4. 设n次多项式有n个根$x_1, \cdots, x_n$，它的判别式是$D=\Pi(x_i-x_j)^2$，这里的乘积是取遍满足$i<j$的所有下标对.

(a) 证明：有理系数多项式的判别式是有理数.

(b) 对于二次多项式，把D明显地表示成多项式系数的有理函数.

*(c) 对三次多项式做同样的问题.

5. 证明：如果K在F上是正规的，并且$F\subset L\subset K$，那么K在L上是正规的.

§ 15.7　子群与子域

如果H是域N的任意自同构集合，那么N中所有在H的全部自同构之下保持不变的元素a（对H中每个T，有$aT=a$）构成N的一个子域. 特别是，如果N是任意多项式在任意基域F上的根域，并且H是N在F上的伽罗瓦群的任意子群，则上述结论也是正确的.

定理 16　*如果H是域N的任意有限自同构群，而K是由所有在H之下不变的元素组成的子域，那么N在K上的次数$[N:K]$至多等于H的阶.*

证明①　如果H的阶是n，那么只须证明，N中任意$n+1$个元

① 这个证明应归于阿廷(Artin)教授. 它包含着这样一个思想：即认为伽罗瓦群只不过是有限自同构群，与基域没有明显的关系.

素 c_1, …, c_{n+1} 在K上是线性相关的. 从H的 n 个元素T出发我们构造 $n+1$ 个未知数 y_i 的 n 个齐次方程的方程组

$$y_1(c_1T)+y_2(c_2T)+\cdots+y_{n+1}(c_{n+1}T)=0.$$

根据 § 2.3 定理 10, 这样的方程组在N中总有一组不同于 $y_1=y_2=\cdots=y_{n+1}=0$ 的解. 现在选取最小的整数 m, 使得 n 个方程

$$y_1(c_1T)+y_2(c_2T)+\cdots+y_m(c_mT)=0,\quad T\in H \tag{14}$$

还有这样的解. 这组解 $y_1,\cdots,y_m$ 是由N的元素组成的,并且除常数因子外,它们是唯一的,这是因为如果有两组不成比例的解,那么通过适当的线性组合将得到含有 $m-1$ 个未知数的方程组的解. 不失一般性,我们也可以假定 $y_1=1$.

现在把H中的任意自同构S作用到(14)式的左边. 因为 $TS=T'$ 跑遍H的所有元素,所以得到方程组

$$(y_1S)(c_1T')+(y_2S)(c_2T')+\cdots+(y_mS)(c_mT')=0,\ T'\in H.$$

除了方程排列的次序外,它与(14)是一样的. 因此 $y_1S,\cdots,y_mS$也是(14)的解,并且根据解的唯一性,这组解就是 $ty_1,\cdots,ty_m$, 其中 t 是比例因子. 然而,因为 $y_1=1$, S 是自同构,所以 $y_1S=1$, 于是 $t=1$. 我们得出结论: 对每个 $i=1,\cdots,m$, 和H中每个 S, 有 $y_iS=y_i$,这就意味着: 系数 y_i 属于H之下不变的元素组成的子域K中. 方程(14)中取 $T=I$, 这就表明元素 $c_1,\cdots,c_m$ 在域K上是线性相关的. 这就证明了定理.

根据这个定理,我们至少可以对可分多项式建立伽罗瓦群的子群与相应根域的子域之间的对应. 这个对应为把已知方程的根域问题化为平行的(有限)伽罗瓦群的子群问题提供了一个系统方法.

定理 17(伽罗瓦理论基本定理) 如果G是F上的可分多项式 $f(x)$的根域N的伽罗瓦群,那么存在G的子群H和N中包含F的子域K之间的双射 $H\longleftrightarrow K$. 如果K已给定,则对应的子群H

$=H(K)$是由G中所有保持K的元素不变的自同构组成；如果H已给定，则对应的子域$K=K(H)$是由N中所有在子群H的每个自同构之下保持不变的元素组成. 对于每个K，子群$H(K)$是N在K上的伽罗瓦群，它的阶等于N在K上的次数$[N:K]$.

证明 对于给定的K，这样来描述$H(K)$:

$$T \text{ 在 } H(K) \text{ 中} \quad \text{当且仅当 } bT=b (\text{对 } K \text{ 中所有 } b). \tag{15}$$

如果S和T具有这种性质，则乘积ST也具有这种性质，所以集合$H(K)$是一个子群. 域N是$f(x)$在K上的根域，N在K上的每个自同构一定是N在F上保持K中每个元素不变的自同构，因此它在子群$H(K)$中. 所以根据定义，$H(K)$是N在K上的伽罗瓦群. 如果把定理 12 应用于这个伽罗瓦群，就可证明$H(K)$的阶数恰好是N在K上的次数.

两个不同的中间域K_1和K_2确定不同的子群$H(K_1)$和$H(K_2)$. 为了证明这一点，选择任意一个在K_1中不在K_2中的元素a，并对N在K_2上的群$H(K_2)$应用定理 13，这就可以断言，$H(K_2)$包含某个T使得$aT\neq a$. 因为a是在K_1中，所以这个自同构T不在群$H(K_1)$中，于是$H(K_1)\neq H(K_2)$.

我们现在知道，对应$K\longmapsto H(K)$是N的所有子域和G的某些子群之间的双射. 为了建立所有子域和所有子群之间的双射，我们必须指出每个子群表现为$H(K)$. 设H为h阶子群，$K=K(H)$象在定理 17 中那样定义：

$$b \text{ 在 } K(H) \text{ 中} \quad \text{当且仅当 } bS=b (\text{对 } H \text{ 中所有 } S). \tag{16}$$

根据定理 16，有$[N:K]\leqslant h$. 比较(15)与(16)我们看出，对应于$K=K(H)$的子群$H(K)$必然包含原来给出的群H，而根据定理 12，$H(K)$的阶是$[N:K]$. 因为$[N:K]\leqslant h$，所以这就意味着群$H(K)$的阶不超过它的子群H的阶. 因此$H(K)=H$，如断言所述. 这就完成了定理的证明.

N 和 F 中间的所有子域 K 组成的集合，对于子域之间的普通包含关系来说，它是一个格．如果 K_1 和 K_2 是两个子域，它们的最大下界(或在这个格中的交)是交集 $K_1 \cap K_2$，它是由 K_1 和 K_2 的所有公共元素组成，而它们的最小上界(或在这个格中的并)是 $K_1 \vee K_2$，它是由 K_1 和 K_2 的全体元素共同生成的 N 的子域．例如，如果 $K_1 = F(v_1)$ 和 $K_2 = F(v_2)$ 都是单扩张，那么它们的并就是多重扩张 $F(v_1, v_2)$．

定理 18 所有子域 $K_1, K_2, \cdots$ 组成的格，通过定理 17 中所述的对应 $K \longmapsto H(K)$，按照下述方式映射到由 G 的所有子群组成的格上：

$$\text{由 } K_1 \subset K_2 \quad \text{可推出 } H(K_1) \supset H(K_2), \tag{17}$$

$$H(K_1 \vee K_2) = H(K_1) \cap H(K_2), \tag{18}$$

$$H(K_1 \cap K_2) = H(K_1) \vee H(K_2). \tag{19}$$

特别是，仅由单位元素组成的子群对应着整个正规域 N．

这些结果表明，这个对应把包含关系颠倒过来，把任意交映射到并，并且把并映射到交．具有这些性质的两个格之间的任意双射称为对偶同构．

为了证明这个定理，我们首先注意，对应于域 K 的群的定义(15)表明，对应于较大子域的群一定使更多的元素保持不变，因此这个群就较小，这就得到(17)．交和并纯粹按照包含关系来定义(见 § 11.7)，因此根据对偶原理，使包含关系颠倒的双射一定把交与并对换，这就是(18)和(19)式所断言的．

我们省略了下述进一步结果的证明．

定理 19 满足 $N \supset K \supset F$ 的域 K 是 F 上的正规域当且仅当它所对应的群 $H(K)$ 是 N 的伽罗瓦群的正规子群．如果 K 是正规的，那么 K 在 F 上的伽罗瓦群 K 与商群 $G/H(K)$ 同构．

这个定理的结论已经在 § 15.4 的末尾解释过了，那里所举的

例子是这个定理的特殊情形.

习　　题

1. (a) 证明：如果H是域N的任意自同构集合，那么N中所有在H的全部自同构之下保持不变的元素构成N的一个子域K.

(b) 证明：N在这个子域K上是正规的.

2. 对$\mathbf{Q}$上的域$\mathbf{Q}(\sqrt{2}, i)$，完整地列出它的子域与子群的对应.

3. 对§15.4中所讨论的x^4-3的根域，做与习题2相同的问题.

4. 证明：$H(K)$在G中的指数等于K在F上的次数.

5. 证明：如果N是F上可分多项式$f(x)$的根域，那么N和F之间的中间域个数是有限的.

6. 证明：N和F之间的所有中间域K构成格.

7. 证明：如果K是特征为∞的域F的有限扩张，那么K和F之间的中间域个数是有限的.

*8. 证明定理19.

*9. 在定理17意义下的两个子域K_1和K_2，如果存在N在F上的一个自同构T把K_1映射到K_2，则称K_1和K_2是共轭的. 证明：K_1和K_2是共轭的当且仅当$T^{-1}H(K_1)T=H(K_2)$（也就是说，当且仅当$H(K_1)$和$H(K_2)$是G的共轭子群）.

§15.8　三次不可约方程

伽罗瓦理论可以用来证明，关于用根式解方程的各种经典问题的不可解性. 作为这个方法的简单例子，我们来考虑有名的具有实根的三次不可约方程.

一个三次方程可以取为如下形式（见§5.5(17)式）

$$f(y)=y^3-py+q=(y-y_1)(y-y_2)(y-y_3), \tag{20}$$

该方程含有实系数p和q，并有三个实的或复的根y_1, y_2, y_3. 系数p和q可以表示成这些根的对称函数，因为当把(20)乘出来时，我们便得到

$$0=y_1+y_2+y_3,\quad p=y_1y_2+y_1y_3+y_2y_3,\quad q=-y_1y_2y_3. \tag{21}$$

引进三次方程的判别式是很重要的，它用下面的公式来定义：

$$D=[(y_1-y_2)(y_1-y_3)(y_2-y_3)]^2. \tag{22}$$

任意两个根的置换不改变D，所以D是y_1, y_2和y_3的对称多项式。根据定理15，可推出D可表示成域$F=\mathbf{Q}(p,q)$中的元素，F是由系数p和q生成的。这个表达式象§5.5(24)中的一样，是

$$D=-4p^3-27q^2. \tag{23}$$

这个等式是y_1, y_2, y_3的多项式恒等式，并可用方程(21)和(22)直接验证。

定理20 具有正判别式的实三次方程有三个实根；如果$D=0$，则至少有两个根是相等的；如果$D<0$，则有两个根是虚根。

只要考察各种类型的根对D的公式(22)有什么影响，就可以验证我们的定理。如果所有的根都是实的，D显然是正的；而如果两个根相等，则$D=0$。最后，假定一个根$y_1=a+bi$是虚数($b\neq 0$)，那么它的复共轭$y_2=a-bi$也一定是一个根(§5.4)，而第三个根是实根。在(22)中，$y_1-y_2=(a+bi)-(a-bi)=2bi$是纯虚数，而

$$(y_1-y_3)(y_2-y_3)=(y_1-y_3)(y_1^*-y_3)=(y_1-y_3)(y_1-y_3)^*$$

是一个实数，所以判别式D就是负的。这恰好给出定理中所列出的几种可能情形。

定理21 如果三次多项式(20)在$F=\mathbf{Q}(p,q)$上是不可约的，有根y_1, y_2, y_3和判别式D，那么它的根域$F(y_1,y_2,y_3)$是$F(\sqrt{D}, y_1)$。

证明 根据D的定义(22)，这个根域一定包含$\sqrt{D}$，因此只剩下证明根y_2, y_3包含在域$K=F(\sqrt{D}, y_1)$中。在域K中，这个三次多项式有一个线性因子$y-y_1$，所以剩下来的二次因子

$$(y-y_2)(y-y_3)=y^2-(y_2+y_3)y+y_2y_3, \tag{24}$$

它的系数还在K中. 把y_1代入(24)中, 则$(y_1-y_2)(y_1-y_3)$在K中, 所以

$$y_2-y_3=\pm\frac{\sqrt{D}}{(y_1-y_2)(y_1-y_3)}$$

在K中. 但是(24)的系数y_2+y_3也在K中. 因为y_2+y_3和y_2-y_3都在K中, 所以y_2和y_3也在K中. 这就证明了定理.

现在考虑在其系数域上是不可约的三次多项式, 它有三个实根. §5.5的公式(19)给出这些根是$y=z-\frac{p}{3z}$, 其中

$$z^3=-\frac{q}{2}+\sqrt{\frac{q^2}{4}+\frac{p^3}{27}}=-\frac{q}{2}+\sqrt{-\frac{D}{108}}.$$

(这里我们用了D的表达式(23).)因为这些根是实的, 所以D是正的(定理20), 因此上面公式中的平方根是虚数. 于是, 这个公式是通过复数给出实根y!

多少年来, 这被认为是这组公式的一个严重缺点, 并有很多数学家尽力寻找求三次方程实根的其他只包含实根式(平方根, 立方根或高次方根)的公式. 但是这种探求都落空了, 这是由于下面定理的缘故.

定理 22 如果一个三次多项式有实根, 并在由它的系数生成的域$F=\mathbf{Q}(p,q)$上是不可约的, 那么不存在求三次多项式实根的有理公式, 这个公式是通过F上的实根式来表示的.

在证明这个定理之前, 我们更普遍地讨论根式$\sqrt[m]{a}=a^{\frac{1}{m}}$的性质. 如果$m$是复合数, 满足$m=rs$, 那么$a^{\frac{1}{m}}=(a^{\frac{1}{r}})^{\frac{1}{s}}$, 等等, 所以任意根式可以通过一系列素指数的根式而得到. 在后面的情形中, 我们可以确定通过添加一个根式所得到的域的次数.

引理 实域[①]K上的r(素数)次多项式x^r-a或者在K上不

① 实域是指其元素为实数的任意域. 这个引理对于任意域都是正确的, 当K的特征是r时证明必须稍微修正一下.

可约，或者在K中有根.

证明 把r次本原单位根ζ添加到K上，然后再把x^r-a的一个根u添加上去. 所得到的扩张$K(\zeta, u)$包含多项式x^r-a的r个根$u, \zeta u, \zeta^2 u, \cdots, \zeta^{r-1}u$，因此它是这个多项式的根域，这个多项式有因子分解式

$$x^r-a=(x-u)(x-\zeta u)(x-\zeta^2 u)\cdots(x-\zeta^{r-1}u).$$

假设x^r-a在K上有正次数$m<r$的真因子$g(x)$. 那么这个因子$g(x)$是x^r-a在$K(\zeta, u)$上的m个线性因子之积，所以$g(x)$中的常数项b是m个根$\zeta^i u$的乘积. 因此$b=\zeta^k u^m$，对某个整数k，并且有

$$b^r=(\zeta^k u^m)^r=(\zeta^r)^k(u^r)^m=(u^r)^m=a^m.$$

由此我们可以在K中求出a的r次根，这是因为$m<r$与r互素，于是存在整数s和t使得$sm+tr=1$(§1.7(13))，所以

$$b^{sr}=a^{sm}=a^{1-tr}=\frac{a}{a^{tr}},$$

则有$a=(b^s a^t)^r$. 于是由x^r-a在K上是可约的这一假定得出x^r-a在K中有一个根$b^s a^t$. 证毕

我们现在可以证明定理22，为此，假定这个结论是错误的. 那么这个三次多项式的某个根可以通过实根式表示，这就是说，根y_1在某个域$L=F(\sqrt[r]{a}, \sqrt[s]{b}, \cdots)$中，域$L$是在$F$上添加实根式而生成的. 因为$D$是正的，把实根式$\sqrt{D}$添加到这个域上得到另一个实域$K=L(\sqrt{D})$. 根据定理21，这三次多项式的全部根都在这个域中，所以它们都可以用含有实根式的公式表示. 域K可以由有限多个根式得到. 如果首先添加$\sqrt{D}$，这就相当于说，K是下面域的有限链中最后一个

$$F\subset K_1\subset K_2\subset K_3\subset\cdots\subset K_n=K, \tag{25}$$

其中

$$K_1=F(\sqrt{D}),\ K_{i+1}=K_i(a_i^{\frac{1}{r_i}}),\quad i=1,\cdots,n-1, \tag{26}$$

这里每个 a_i 在 K_i 中，每个 r_i 是素数. 去掉额外的域，我们可以假定实根 $a_i^{\frac{1}{r_i}}$ 不在域 K_i 中，根据引理，这就意味着 $x^{r_i}-a_i$ 在 K_i 上是不可约的，因此 K_{i+1} 的次数是 $[K_{i+1}:K_i]=r_i$.

根据假定，三次多项式的根在 K 中，它们并不在 F 中或者 $F(\sqrt{D})$ 中，这因为三次多项式在 F 上是不可约的. 那么在链(25)中存在第一个包含三次多项式的一个根(比如说 y_1)的域 K_{j+1}. 在前一个域 K_j 上，已知的三次多项式一定是不可约的，如果不然，它在 K_j 上将有线性因子 $y-y_i$，这与 K_j 不包含任何一个 y_i 这一事实相矛盾. 那么扩张

$$K_{j+1}=K_j(a^{\frac{1}{r}}),\quad a=a_j,\quad r=r_j \tag{27}$$

的次数是 r，并且包含一个元素 y_1，y_1 在 K_j 上的次数是 3. 根据 §14.5 定理 9 的推论 2，有 $3|r$，所以素数 r 一定是 3，于是我们在(27)中讨论三次根 $\sqrt[3]{a}$. 这个域 K_{j+1} 是在 K_j 上添加 y_1 而生成的，它包含 $\sqrt{D}$，因此根据定理 21，它包含这个三次多项式的全部根. 所以 K_{j+1} 是给定的三次多项式在 K_j 上的根域. 根据定理 14，它作为一个根域，是正规的. 因为它包含 K_j 上不可约多项式 x^3-a 的一个根 $a^{\frac{1}{3}}$，所以它一定包含这个方程的全部根. 其他两个根是 $\omega a^{\frac{1}{3}}$ 和 $\omega^2 a^{\frac{1}{3}}$，所以 K_{j+1} 也包含三次复单位根 ω. 这与 $K_{j+1}\subset K$ 是实域的假定相违背. 定理证完.

习　题

1. 验证判别式(23).

2. 依据定理 21 的方法，用 y_1 和 $\sqrt{D}$ 明显地表示出三次多项式的所有根.

*3. 三次方程的讨论，有哪些可以应用到 $\mathbf{Z}_3$ 上的三次方程.

4. 证明：一个多项式 x^n-a，如果在特征为∞的域 F 上有次数与 n 互素的因子，那么它在 F 中有根.

5. 证明：如果 F 是特征为∞的域，它包含所有 n 次单位根，那么次数 $[F(a^{\frac{1}{n}}):F]$ 是 n 的因子.

6. 考虑三次不可约多项式(20)在 $F=\mathbf{Q}(p, q)$ 上的伽罗瓦群 G. 证明：如果 D 是 F 的一个数的平方，那么 G 是三个字母的交错群，否则它是对称群.

§15.9 五次方程的不可解性

在本节中，F 表示复数域中包含所有单位根的一个子域，K 表示 F 的各种有限扩张.

假设 $K=F(a^{\frac{1}{r}})$ 是由 F 和 $a\in F$ 的单个 r 次根 $a^{\frac{1}{r}}$ 生成的，这里 r 是素数. 象在第五章那样，$x^r=a$ 的其他根是 $\zeta a^{\frac{1}{r}}, \cdots, \zeta^{r-1}a^{\frac{1}{r}}$，这里 ζ 是 r 次本原单位根，因此它也在 F 中. 根据 §15.8 引理 2，除非 $K=F$，多项式 x^r-a 在 F 上是不可约的，所以有一个 K 的自同构 S 把根 $a^{\frac{1}{r}}$ 映射到根 $\zeta a^{\frac{1}{r}}$. 这个自同构的幂 $I, S, S^2, \cdots, S^{r-1}$ 把 $a^{\frac{1}{r}}$ 分别映射到方程 $x^r=a$ 的每一个根，因此这些幂包含 K 在 F 上的全部自同构. 于是我们得出结论：*K 在 F 上的伽罗瓦群是循环群.*

更一般地，假设 K 在 F 上是正规的，并可由 F 通过一系列单扩张得到，每个单扩张只是在前一个 F 的扩张上添加一个 n_i 次根得到的. 这就意味着，存在一系列中间域 K_i

$$F=K_0\subset K_1\subset K_2\subset\cdots\subset K_s=K \tag{28}$$

满足 $K_i=K_{i-1}(x_i)$，这里 $x_i^{n_i}\in K_{i-1}$. 不失一般性，我们可以假定每个 n_i 是素数. 我们把这样的 K 称为 F 的*根式扩张*. 因为 K 是正规的，它是多项式 $f(x)$ 在 F 上的根域，所以它也是同一个多项

式 $f(x)$ 在 K_1 上的根域——所以(根据定理 14) 它在 K_1 上也是正规的. 但是根据上一节, K_1 在 F 上是正规的. 所以 K 在 F 上的每个自同构诱导出 K_1 在 F 上的一个自同构, 并且自同构的乘法也相同. 进一步, 根据 § 15.2 引理 2, K_1 在 F 上的每个自同构可以扩张成 K 在 F 上的一个自同构. 因此对应是从 K 在 F 上的伽罗瓦群到 K_1 在 F 上的伽罗瓦群的满同态, 这很象 § 15.4 末尾所描述的那样. 此外, 在这个满同态之下, 诱导出 K_1 在 F 上的恒等自同构的元素刚好是由 K 在 K_1 上的自同构定义的. 这就表明, 在这个满同态之下, K 在 F 上的伽罗瓦群 $G(K/F)$ 被映射到 $G(K_1/F)$ 上. 因此 $G(K_1/F)$ 与商群 $G(K/F)/G(K/K_1)$ 同构. 把这个结果同上一节的结果结合起来, 我们推出 $G(K/K_1)$ 是 $G(K/F)$ 的正规子群, 并且商群 $G(K_1/F)$ 是循环商群.

现在对 s 用归纳法. 根据定义, K 是 K_1 的根式扩张; 如上所述, 它在 K_1 上也是正规的. 因此前面的推理可以再应用到 $G(K/K_1)$ 上, 证明 $G(K/K_2)$ 是 $G(K/K_1)$ 的正规子群, 它们构成循环商群 $G(K_2/K_1)$. 把这个论证重复 s 次, 并用 S_i 表示子群 $G(K/K_i)$, 我们就得出下面基本结果.

定理 23 *设 K 是 F 的任意正规根式扩张, 那么 K 在 F 上的伽罗瓦群 G 包含一个子群序列 $S_0=G\supset S_1\supset S_2\supset\cdots\supset S_s$, 其中每个子群是它前一子群的正规子群, 并且商群 S_{i-1}/S_i 是循环商群, 而 S_s 仅由 I 组成.*

这表明 K 在 F 上的伽罗瓦群 G 在下面定义之下是可解的.

定义 *有限群 G 是可解的当且仅当它包含一个子群链 $S_0=G\supset S_1\supset S_2\supset\cdots\supset S_s=I$ 使得对所有的 k, 有 (i) S_k 是 S_{k-1} 中的正规子群; (ii) S_{k-1}/S_k 是循环群.*

关于抽象可解群, 有很大一部分是知道的, 例如, 任意一个群, 如果它的阶可被少于三个的不同素数整除, 那么它是可解群(波恩

赛德(Burnside)). 甚至知道，每个奇数阶群是可解的（费特-汤姆森(Feit-Thompson)). 然而，我们只满足于下面的简单事实.

引理1 *有限可解群G的任意满同态象G'本身是可解的.*

证明 设G有一个象可解性定义中所描述的那样的子群链$S_0=G\supset S_1\supset S_2\supset\cdots\supset S_s=I$, 并设$S'_0=G', S'_1,\cdots,S'_s=I'$是它们的满同态象. 那么，每个$S'_k$，如果它包含$x'$和$y'$，则必包含$x'y'=(xy)'$和$(x')^{-1}=(x^{-1})'$($x, y$分别是$x'$和$y'$在$S_k$中的象源)，所以$S'_k$是$G'$的子群. 进一步，如果$a$在$S_{k-1}$中，$x$在$S_k$中，那么$S_k$在$S_{k-1}$中的正规性意味着$a^{-1}xa$在$S_{k-1}$中，因此$(a')^{-1}x'a'=(a^{-1}xa)'$在$S'_k$中. 因为$a'$可以是$S'_{k-1}$的任意元素，所以这就证明了$S'_k$在$S'_{k-1}$中是正规的. 最后，因为$S_{k-1}$是由$S_k$的某单个陪集的各次幂$(S_ka)^n=S_ka^n$组成($S_{k-1}/S_k$是循环的)，所以$S'_{k-1}$是这个陪集的象的各次幂$S'_k(a')^n=(S'_ka')^n$组成，于是$S'_{k-1}/S'_k$是循环的. 所以这些子群组成的链$S'_0\supset S'_1\supset S'_2\supset\cdots\supset S'_s=I'$具有使$G'$是可解群的全部性质，正如引理1所要求的. 证毕

现在让我们来定义，系数在F中的一个方程$f(x)=0$，如果它的根在F的扩张K中，这个K可以通过在F上逐次添加n次根而得到，我们就称$f(x)=0$在F上是*根式可解的*. 根据§5.5，所有二次方程、三次方程和四次方程，情况都是如此. 应看到，K并不要求是正规的，而只要求包含$f(x)$在F上的根域N. 然而，因为可用根式表示的元素的任意共轭元素本身可用共轭根式表示，所以$f(x)$的根域N也一定包含在有限扩张$K^*\supset K$中，K^*在F上是正规的，并是F的根式扩张. 这个K^*包含N作为F上的正规子域，因此K^*在F上的每个自同构S诱导出N在F上的一个自同构S_1，而且对应$S\longmapsto S_1$是满同态. 这就是说，K^*在F上的伽罗瓦群与N在F上的伽罗瓦群是满同态的，但是，前者是可解的(根据定理23)，因此根据引理1，后者也是可解的. 这就证明了

定理 24 如果系数在 F 中的一个方程 $f(x)=0$ 是根式可解的，那么它在 F 上的伽罗瓦群是可解的。

为了证明五次方程不总是根式可解的，我们只需要找出一个方程，它的伽罗瓦群不是可解的．我们就来做这件事．首先我们证明五次对称群不是可解的，然后再列出一个五次方程，它的伽罗瓦群是五次对称群．

定理 25 n 个字母的对称群，除 $n\leqslant 4$ 之外，不是可解的。

证明 设 $G=S_0\supset S_1\supset S_2\supset\cdots\supset S_s$ 是任意子群链，其中每个子群在前一个子群中是正规的，并且商群 S_{s-1}/S_s 是循环商群，我们通过对 s 用归纳法来证明，S_s 一定包含每个三字母循环 $(i\,jk)$。由此推出 $S_s>I$，所以 G 不能是可解的．

因为 $S_0=G$ 包含每个三字母循环，所以我们只须用归纳法证明，如果 S_{s-1} 包含每个三字母循环，那么 S_s 也包含每个三字母循环．首先注意，如果置换 ϕ 和 ψ 两个都在 S_{s-1} 中，那么它们的所谓"换位子"$\gamma=\phi^{-1}\psi^{-1}\phi\psi$ 在 S_s 中．为了看出这一点，考虑 S_{s-1}/S_s 中的象 ϕ', ψ' 和 γ'．这个商群是循环的，也是交换群，因此在 S_{s-1}/S_s 中有

$$\gamma'=(\phi')^{-1}(\psi')^{-1}\phi'\psi'=I',$$

这意味着 $\gamma\in S_s$．但是在 $\phi=(i\,l\,j)$ 和 $\psi=(jkm)$ 这一特殊情形中，这里 i, j, k 是给定的，l, m 是任意两个其他字母(除 $n\leqslant 4$ 之外，这样的字母是存在的)，我们有

$$\gamma=(j\,l\,i)(mkj)(i\,l\,j)(jkm)=(i\,jk)\in S_s,\text{对所有的 } i, j, k.$$

这就证明了 S_s 包含每个三字母循环，这就是所要求的．

附带说一下，证明这个定理的更明显的形式是可能的．我们知道，交错群 A_n 是对称群 G 的正规子群，所以存在一个始于 $G\supset A_n$ 的群链．然后我们可以证明交错群 A_n(当 $n>4$)，除它本身和单位元素之外没有任何正规子群．

引理 2 存在一个(实)五次方程，它的伽罗瓦群是五个字母的对称群.

证明 设 A 是所有代数数构成的域，它是可数的，并且包含所有单位根. 因此，我们可以象 § 14.6 那样相继选取五个在 A 上代数无关的实数 $x_1, \cdots, x_5$，构成超越扩张 $A(x_1, \cdots, x_5)$. 现在设 $\sigma_1, \cdots, \sigma_5$ 是 $x_1, \cdots, x_5$ 的初等对称多项式，并设 $F=A(\sigma_1, \cdots, \sigma_5)$. 象在定理 15 中那样，多项式

$$f(t)=t^5-\sigma_1 t^4+\sigma_2 t^3-\sigma_3 t^2+\sigma_4 t-\sigma_5 \qquad (29)$$

在 F 上的伽罗瓦群是五个字母 $x_1, \cdots, x_5$ 的对称群.

由引理 2 和定理 25 推出，在含有一切单位根的域上存在一个(实)五次方程，它的伽罗瓦群不是可解的. 现在应用定理 24，我们得到我们的最后结论

定理 26 存在(实)五次方程，不能用根式求解.

习　　题

1. 证明：三个字母的对称群是可解的.

2. 证明：任意有限交换群是可解的. (提示：证明它包含一个素指数的(正规)子群.)

3. 证明：如果有限群 G 包含一个正规子群 N，使得 N 和 G/N 都是可解的，那么 G 是可解的.

4. (a) 证明：在四个字母的对称群中，三字母循环的全体换位子构成四阶正规子群.

(b) 利用(a)和交错子群来证明：四个字母的对称群是可解的.

5. 证明：任意有限抽象群 G 是某适当方程的伽罗瓦群. (提示：根据凯莱定理，G 与对称群的子群同构.)

*6. (a) 证明：$x^n=a$ 的伽罗瓦群甚至在不包含单位根的域上也是可解的.

(b) 证明：定理 24 对于任意域 F，不管它是否包含单位根，都成立.

7. 明显地证明：如果K是F的根式扩张，那么存在K的一个扩张K^*，它在F上是正规的，并且它也是F的根式扩张。（上面定理24的证明中用到了这个事实。）

8. 证明：如果F包含n次单位根，$K=F(a^{\frac{1}{n}})$，这里a在K中，那么，甚至当n不是素数时，K在F上的伽罗瓦群也是循环的。

9. 证明：如果$\mathbf{Q}$是有理数域，f是(29)式表示的特殊多项式，那么f在域$\mathbf{Q}(\sigma_1, \cdots, \sigma_5)$上的伽罗瓦群仍然是五个字母的对称群。

10. 证明：如果$n>4$，那么存在一个实n次方程不是根式可解的。

文献目录

一般参考文献

Albert, A. A. (ed.). *Studies in Modern Algebra* (MAA Studies in Mathematics, II). Englewood Cliffs, N.J.: Prentice-Hall, 1963.

Artin, E. *Geometric Algebra.* New York: Interscience, 1957.

Birkhoff, G., and T. C. Bartee. *Modern Applied Algebra.* New York: McGraw-Hill, 1970.

Godement, Roger. *Cours d'algébre.* Paris: Hermann, 1963.

Herstein, I. N. *Topics in Algébra.* New York: Wiley, 1964.

Jacobson, N. *Basic Algebra. I. Basic Algebra. II.* San Francisco: Freeman, 1974, 1976.

Mac Lane, Saunders, and Garrett Birkhoff. *Algebra.* New York, Macmillan, 1967.

Schreier, O., and E. Sperner. *Introduction to Modern Algebra and Matrix Theory* (English translation). New York: Chelsea, 1952.

Uspensky, J. V. *Theory of Equations.* New York: McGraw-Hill, 1948.

van der Waerden, B. L. *Modern Algebra,* I, 4th ed., and II, 5th ed. (English translation). New York: Ungar, 1966 and 1967.

数论

Hardy, G. H., and E. M. Wright. *An Introduction to the Theory of Numbers*, 4th ed. Oxford: Clarendon, 1954.

LeVeque, W. J. *Topics in Number Theory.* 2 vols. Reading, Mass.: Addison-Wesley, 1956.

Niven, Ivan, and H. S. Zuckerman. *An Introduction to the Theory of Numbers.* New York: Wiley, 1960.

Rademacher, H. *Lectures on Elementary Number Theory.* New York: Wiley, 1964.

代数数论

Lang, S. *Algebraic Numbers.* Reading, Mass.: Addison-Wesley, 1964.

Ribenboim, P. *Algebraic Numbers.* New York: Wiley, 1972.

Weiss, E. *Algebraic Number Theory.* New York: McGraw-Hill, 1963.

群论

Curtis, C. W., and I. Reiner. *Representation Theory of Finite Groups and Associative Algebras.* New York: Interscience, 1962.

Fuchs. L. *Abelian Groups.* Budapest: Hungarian Academy of Sciences, 1958.

Gorenstein, D. *Finite Groups.* New York: Harper & Row, 1968.

Hall, M. *The Theory of Groups.* New York: Macmillan, 1959.

Rotman, J. J. *The Theory of Groups.* Boston: Allyn & Bacon, 1965.

矩阵论

Faddaeva, V. N. *Computational Methods of Linear Algebra.* Translated by C. D. Benster. New York: Dover, 1959.

Varga, R. S. *Matrix Iterative Analysis.* Englewood Cliffs, N. J.: Prentice-Hall, 1962.

伽罗瓦理论

Artin, E. *Galois Theory*, 2nd ed. (Notre Dame Mathematical Lecture No. 2). Notre Dame, Ind.: University of Notre Dame Press, 1944.

线性代数与环

Jacobson, N. *Lie Algebras.* New York: Wiley, 1962.

Jacobson, N. *The Structure of Rings*, 2nd ed. New York: American Mathematical Society, 1964.

McCoy, N. H. *The Theory of Rings.* New York: Macmillan, 1964.

代数几何

Fulton, W. *Algebraic Curves.* New York: Benjamin, 1969.

Jenner, W. E. *Rudiments of Algebraic Geometry.* New York: Oxford University Press, 1963.

Lang, S. *Introduction to Algebraic Geometry.* New York: Interscience, 1958.

Zariski, O., and P. Samuel. *Commutative Algebras.* 2 vols. New York: Van Nostrand, 1958, 1960.

逻辑学

Kleene, S. C. *Mathematical Logic*. New York: Wiley, 1967.

Mendelson, E. *Introduction to Mathematical Logic*. New York: Van Nostrand, 1964.

格论

Abbott, J. C. *Sets, Lattices and Boolean Algebras*. Boston: Allyn & Bacon, 1969.

Birkhoff, Garrett. *Lattice Theory*, 3rd ed. Providence: American Mathematical Society, 1966.

同调代数

Freyd, P. *Abelian Categories*. New York: Harper & Row, 1964.

Jans, J. P. *Rings and Homology*. New York: Holt, 1964.

Mac Lane, Saunders. *Homology*. Berlin: Springer, 1963.

Mac Lane, Saunders. *Categories for the Working Mathematician*. Berlin: Springer, 1971.

泛代数

Cohn, P. M. *Universal Algebra*. New York: Harper & Row, 1965.

Grätzer, G. *Universal Algebra*. New York: Van Nostrand, 1968.

Jonsson, Bjarni. *Topics in Universal Algebra* (Lecture Notes in Mathematics No. 250). Berlin: Springer, 1972.

数学符号表

符号	含义
A	矩阵（B, C 等也是）
A'	转置矩阵
A^*	矩阵的复共轭
$\mathfrak{A}$	线性代数
$A_n(F)$	F 上仿射群
B	布尔代数
c	**R** 的基数
C	复数域
D	整环
$D[x]$	系数在 D 中的 x 的多项式形式
$D\langle x\rangle$	系数在 D 中的 x 的多项式函数
d	正整数的集合的基数 $o(\mathbf{Z}^+)$
E_n	n 维欧几里得空间
E_{ij}	特殊矩阵，(i, j)位置的元素为 1，其他位置为 0
e, 1	群的单位元素
F	域
F^n	F 上 n-数组组成的空间
$F[x]$	系数在 F 中的 x 的多项式形式
$F(x)$	系数在 F 中的 x 的有理形式
G	群
g.l.b.	最大下界
i	$\sqrt{-1}$；四元数单位
I	恒等变换或单位矩阵；格的最大元素
j, k	四元数单位
J	环中的理想(H, L等等也是)
K	域
$[K:F]$	K 在 F 上的次数
$L_n(F)$	F 上全线性群
l.u.b.	最小上界
$M_n(F)$	F 上全阵代数
O	零矩阵；格的最小元素
$O_n(F)$	正交群
$o(S)$	集合 S 的基数
P	素理想；非奇异矩阵

p, q	正素数
$Q(D)$	整环D的商域
$\mathbf{Q}$	有理数域
R	环
$\mathbf{R}$	实数域
S	集合；子群；子空间
S'	集合S的补
$S^{\perp}$	子空间的正交补
T	线性变换
T_A	用矩阵A给出的线性变换
$[u:F]$	u在F上的次数
V, W	矢量空间
V^*	对偶矢量空间
X	矢量或行矩阵
z^*	共轭复数
$\mathbf{Z}$	整数环或整数群
$\mathbf{Z}_n$	模n整数环
$\mathbf{Z}^+$	正整数集合
α, β	矢量
(α, β)	矢量内积(点积)
$\alpha \times \beta$	矢量外积(矢量积)
δ_{ij}	克罗内克尔符号
ε_i	单位矢量
ϕ, ψ	变换；映射；函数
$\prod$	乘积
$\sum$	求和
ξ, η	矢量
$\mathbf{0}$	零矢量
$\emptyset$	空集
$\cap, \cup$	交，并(集合的)
$\wedge, \vee$	交，并(布尔代数，格中的)
$\in$	属于；是…的元素
$\subset$	包含于；是…的子集合
$<$	小于；真包含在…中
$\leqslant$	不等号
$\perp$	正交于；垂直于
$\otimes$	张量积
$\times$	直积
$\oplus$	直和
$\circ$	二元运算
$\longmapsto$	元素的映入
$\longrightarrow$	集合的映入
∞	无限；无穷大
$\sim$	相伴
$\equiv$	同余
$\lvert A\rvert$	矩阵A的行列式，也记作 $\det A$
$\lvert a\rvert$	绝对值
(a_{ij})	矩阵
$a \mid b$	a整除b
(a, b)	最大公因子(g. c. d.)
$[a, b]$	最小公倍数(l. c. m.)

索　　引

四　画

五　画

六　画

七　画

八　画

九　画

十　画

十 一 画

十 二 画

十三画以上